

DÉPARTEMENT DE GÉOMATIQUE APPLIQUÉE  
Faculté des lettres et sciences humaines  
Université de Sherbrooke



Modélisation des densités moyennes annuelles de populations de moustiques associées à des  
maladies zoonotiques vectorielles représentant un risque pour la santé publique  
pour le sud du Québec

par  
Julie ALLOSTRY

Mémoire présenté pour l'obtention du grade de Maître ès sciences géographiques (M.Sc.),  
cheminement Géomatique

Avril 2019

© Julie Allostry, 2019

## Table des matières

Remerciements .....	viii
Description du projet de recherche.....	ix
Avant-propos .....	x
Résumé .....	xi
1. Introduction .....	1
2. Hypothèse et objectifs .....	8
3. Matériel .....	9
3.1. Échelle de travail .....	9
3.2. Échantillonnage des moustiques.....	9
3.3. Variables environnementales.....	10
3.3.1. Données d'occupation et utilisation du sol .....	10
3.3.2. Données météorologiques .....	11
4. Méthode.....	13
4.1. Schéma méthodologique.....	13
4.2. Phase exploratoire.....	16
4.2.1. Étape 1 : cartographie de la répartition des pièges et analyses descriptives .....	16
4.2.2. Étape 2 : préparation et extraction des variables environnementales.....	16
4.3. Phase de modélisation .....	20
4.3.1. Étape 1 : transformation des données.....	21
4.3.2. Étape 2 : régressions univariées .....	21
4.3.3. Étape 3 : indépendance des variables .....	22
4.3.4. Étape 3 : régressions multivariées.....	22
4.3.5. Étape 5 : choix du meilleur modèle.....	25
4.4. Phase de validation .....	25
4.4.1. Étape 1 : validation interne.....	25
4.4.2. Étape 2 : validation externe .....	26
4.5. Phase cartographique.....	27
4.5.1. Étape 1 : carte des résidus des modèles.....	27
4.5.2. Étape 2 : cartes annuelles des densités d'espèces dans la zone d'étude.....	27
5. Résultats .....	29
5.1. Phase exploratoire.....	29
5.1.1. Étape 1 : cartographie de la répartition des pièges et analyses descriptives .....	29

5.1.2.	Étape 2 : variables environnementales .....	32
5.1.2.1.	Variables d'occupation du sol.....	32
5.1.2.2.	Variables météorologiques.....	32
5.2.	Phase de modélisation .....	34
5.2.1.	Étape 1 : transformation des données.....	34
5.2.2.	Étape 2 : régressions univariées .....	35
5.2.3.	Étape 3 : indépendance des variables .....	36
5.2.4.	Étape 4 : régressions multivariées .....	39
5.2.5.	Étape 5 : choix du meilleur modèle.....	48
5.3.	Phase de validation .....	52
5.3.1.	Étape 1 : validation interne.....	52
5.3.2.	Étape 2 : validation externe .....	62
5.5.	Phase cartographique.....	66
5.5.1.	Étape 1 : carte des résidus des modèles.....	66
5.5.2.	Étape 2 : cartes annuelles des densités d'espèces dans la zone d'étude.....	74
6.	Discussion .....	82
7.	Conclusion.....	87
8.	Références .....	89
ANNEXE	.....	94
	Note sur l'autocorrélation spatiale .....	95
	Note sur l'étude de dispersion des espèces au fil du temps .....	95

# Liste des figures

Figure 1 : Schéma méthodologique.....	15
Figure 2 : Assemblage des données d'occupation du sol de base, agrandissement sur Montréal et ses alentours, 2011. ....	18
Figure 3 : Image des précipitations moyennes du 7 janvier 2003 avec contour des régions administratives. ....	20
Figures 4A et B : Histogramme de la période d'opération des pièges de type <i>Light trap</i> de 2003	29
Figure 5 : Délimitation de la zone d'étude. ....	31
Figures 6A et B : Répartition des pièges selon leur durée de fonctionnement pour 2004 et 2014. ....	31
Figure 7 : Couche des classes d'occupation du sol bonifiée de 2011 – comparaison avec une image satellite de Google Earth – agrandissement sur l'île de Montréal. ....	33
Figure 8 : Relation non linéaire entre la variable centrée réduite Agri et les quantités transformées de <i>Culex pipiens-restuans</i> . ....	44
Figure 9 : Relation non linéaire entre la variable centrée réduite Boise et les quantités transformées de moustiques d' <i>Ochlerotatus stimulans</i> . ....	44
Figure 10 : Relation non linéaire entre la variable centrée réduite Agri et les quantités transformées d' <i>Aedes vexans</i> . ....	45
Figure 11 : <i>P-P plot</i> et <i>Q-Q plot</i> des résidus du meilleur modèle du groupe <i>Culex pipiens-restuans</i> . ....	53
Figure 12 : Répartition des résidus de Student en fonction des prédictions du meilleur modèle du groupe <i>Culex pipiens-restuans</i> . ....	53
Figure 13 : <i>P-P plot</i> et <i>Q-Q plot</i> des résidus du meilleur modèle pour <i>Aedes vexans</i> . ....	54
Figure 14 : Répartition des résidus de Student en fonction des prédictions du meilleur modèle pour <i>Coquillettidia perturbans</i> . ....	54
Figure 15 : <i>P-P plot</i> et <i>Q-Q plot</i> des résidus du meilleur modèle du groupe <i>Ochlerotatus stimulans</i> . ....	55
Figure 16 : Répartition des résidus de Student en fonction des prédictions du meilleur modèle du groupe <i>Ochlerotatus stimulans</i> . ....	55
Figure 17 : <i>P-P plot</i> et <i>Q-Q plot</i> des résidus du meilleur modèle pour <i>Aedes vexans</i> . ....	56
Figure 18 : Répartition des résidus de Student en fonction des prédictions du meilleur modèle pour <i>Aedes vexans</i> . ....	56
Figure 19 : Histogramme de distribution des résidus de Student du MLM1f pour le groupe <i>Culex pipiens-restuans</i> et statistiques descriptives associées (données d'entraînement). ....	58
Figure 20 : Histogramme de distribution des résidus de Student du MLM1f pour <i>Coquillettidia perturbans</i> et statistiques descriptives associées (données d'entraînement). ....	59
Figure 21 : Histogrammes de distribution des résidus de Student du MLM1f pour le groupe <i>Ochlerotatus stimulans</i> et statistiques descriptives associées (données d'entraînement). ....	59
Figure 22 : Histogramme de distribution des résidus de Student du MLM1f pour <i>Aedes vexans</i> et statistiques descriptives associées (données d'entraînement). ....	60
Figures 23A à D : Distances de Cook des résidus de Student pour le meilleur modèle pour le groupe <i>Culex pipiens-restuans</i> (A), <i>Coquillettidia perturbans</i> (B), le groupe <i>Ochlerotatus stimulans</i> (C) et <i>Aedes vexans</i> (D). ....	61
Figure 24 : Histogramme de distribution des résidus de Student du MLM1f pour le groupe <i>Culex pipiens-restuans</i> appliqué aux données de validation et statistiques descriptives associées. ....	63

Figure 25 : Histogramme de distribution des résidus de Student du MLM1f pour <i>Coquillettidia perturbans</i> appliqué aux données de validation et statistiques descriptives associées. ....	63
Figure 26 : Histogramme de distribution des résidus de Student du MLM1f pour le groupe <i>Ochlerotatus stimulans</i> appliqué aux données de validation et statistiques descriptives associées. ....	64
Figure 27 : Histogramme de distribution des résidus de Student du MLM1f pour <i>Aedes vexans</i> appliqué aux données de validation et statistiques descriptives associées. ....	64
Figures 28A et B : Distribution spatiale des résidus de Student du MLM1f du groupe <i>Culex pipiens-restuans</i> pour les données d'entraînement (A) et superposant les données d'entraînement et de validation (B). ....	70
Figures 29A et B : Distribution spatiale des résidus de Student du MLM1f de <i>Coquillettidia perturbans</i> pour les données d'entraînement (A) et superposant les données d'entraînement et de validation (B). ....	71
Figures 30A et B : Distribution spatiale des résidus de Student du MLM1f du groupe <i>Ochlerotatus stimulans</i> pour les données d'entraînement (A) et superposant les données d'entraînement et de validation (B). ....	72
Figures 31A et B : Distribution spatiale des résidus de Student du MLM1f d' <i>Aedes vexans</i> pour les données d'entraînement (A) et superposant les données d'entraînement et de validation (B). ....	73
Figures 32A, B et C : Image d'origine de la classe Urbain (A), image moyennée de la classe Urbain (B) et image centrée réduite de la classe Urbain (C). ....	74
Figure 33 : Application du modèle final du groupe <i>Culex pipiens-restuans</i> aux données environnementales de 2014. ....	75
Figure 34A et B : Cartographie annuelle prédictive du groupe <i>Culex pipiens-restuans</i> validée avec les données de 2014 sur la zone d'étude (A) et agrandie sur le grand Montréal (B). ....	78
Figure 35A et B : Cartographie annuelle prédictive de <i>Coquillettidia perturbans</i> validée avec les données de 2014 sur la zone d'étude (A) et agrandie sur le grand Montréal (B). ....	79
Figure 36A et B : Cartographie annuelle prédictive du groupe <i>Ochlerotatus stimulans</i> validée avec les données de 2014 sur la zone d'étude (A) et agrandie sur le grand Montréal (B). ....	80
Figure 37A et B : Cartographie annuelle prédictive d' <i>Aedes vexans</i> validée avec les données de 2014 sur la zone d'étude (A) et agrandie sur le grand Montréal (B). ....	81

## Liste des tableaux

Tableau 1 : Arbovirus endémiques en Amérique du Nord par potentiel vectoriel des espèces de moustiques ciblées pour le projet. ....	4
Tableau 2 : Répartition temporelle des pièges en fonction de leur source et du type de piège. ....	10
Tableau 3 : Métadonnées. ....	12
Tableau 4 : Règles de classification pour bonifier la couche de base d'Agriculture et Agroalimentaire Canada. ....	17
Tableau 5 : Extrait des pourcentages d'occupation du sol sous les zones tampons de 1 km autour des pièges. ....	32
Tableau 6 : Paramètres de la transformation Box Cox des données de capture. ....	34
Tableau 7 : Test de normalité des données d'entraînement transformées Box Cox. ....	34
Tableau 8 : Statistiques descriptives des données d'entraînement d'origine. ....	35
Tableau 9 : Statistiques descriptives des données d'entraînement transformées Box Cox. ....	35

Tableau 10 : Nombre de variables significatives par espèce au terme des régressions univariées.	36
Tableau 11 : Variables centrées réduites significatives par espèce au terme du test de similarité à 80 %.	37
Tableau 12 : Variables centrées réduites explicatives significatives et indépendantes par espèce au terme du test de multicollinéarité avec le sens de corrélation, positif ou négatif.	37
Tableau 13 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance pour le groupe <i>Culex pipiens-restuans</i> .	39
Tableau 14 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 10 de variation d'inflation de la variance pour le groupe <i>Culex pipiens-restuans</i> .	39
Tableau 15 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance pour <i>Coquillettidia perturbans</i> .	40
Tableau 16 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 5 de variation d'inflation de la variance pour <i>Coquillettidia perturbans</i> .	40
Tableau 17 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 10 de variation d'inflation de la variance pour <i>Coquillettidia perturbans</i> .	40
Tableau 18 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance pour le groupe <i>Ochlerotatus stimulans</i> .	41
Tableau 19 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 5 de variation d'inflation de la variance pour le groupe <i>Ochlerotatus stimulans</i> .	41
Tableau 20 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 10 de variation d'inflation de la variance pour le groupe <i>Ochlerotatus stimulans</i> .	41
Tableau 21 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance pour <i>Aedes vexans</i> .	42
Tableau 22 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance pour <i>Aedes vexans</i> .	42
Tableau 23 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 10 de variation d'inflation de la variance pour <i>Aedes vexans</i> .	42
Tableau 24 : Variables simples et termes quadratiques et/ou d'interactions significatifs pour chaque espèce.	45
Tableau 25 : MLM1avec10 pour le groupe <i>Culex pipiens-restuans</i> .	46
Tableau 26 : MLM1avec3 pour le groupe <i>Ochlerotatus stimulans</i> .	46
Tableau 27 : MLM1avec5 pour le groupe <i>Ochlerotatus stimulans</i> .	47
Tableau 28 : MLM1avec10 pour le groupe <i>Ochlerotatus stimulans</i> .	47
Tableau 29 : MLM1avec10 pour <i>Aedes vexans</i> .	47
Tableau 30 : Évaluation des modèles pour le groupe <i>Culex pipiens-restuans</i> .	48

Tableau 31 : Évaluation des modèles pour <i>Coquillettidia perturbans</i> .....	49
Tableau 32 : Évaluation des modèles pour le groupe <i>Ochlerotatus stimulans</i> . ....	50
Tableau 33 : Évaluation des modèles pour <i>Aedes vexans</i> . ....	51
Tableau 34 : Comparaison des valeurs du MLM1f du groupe <i>Culex pipiens-restuans</i> pour les pièges du secteur de la réserve faunique de Marguerite d'Youville par rapport à la moyenne de tous les pièges.....	68
Tableau 35 : Comparaison des valeurs du MLM1f de <i>Coquillettidia perturbans</i> pour les pièges du secteur de la réserve faunique de Marguerite d'Youville par rapport à la moyenne de tous les pièges.....	69
Tableau 36 : Comparaison des valeurs du MLM1f du groupe <i>Ochlerotatus stimulans</i> pour les pièges du secteur de la réserve faunique de Marguerite d'Youville par rapport à la moyenne de tous les pièges.....	69
Tableau 37 : Comparaison des valeurs du MLM1f d' <i>Aedes vexans</i> pour les pièges du secteur de la réserve faunique de Marguerite d'Youville par rapport à la moyenne de tous les pièges. ....	69
Tableau 38 : Moyennes et écart-types des variables simple d'entraînement appliqués pour centrer réduire les variables de validation correspondantes.....	74
Tableau 39 : Moyennes et écart-types des variables composées d'entraînement appliqués pour centrer réduire les variables de validation correspondantes.....	74

## Abréviations

_1	Variable météorologique de l'année précédant l'année de capture des moustiques
AAC	Agriculture et agroalimentaire Canada
AIC	<i>Akaike Information Criterion</i> (Critère d'information Akaike)
ASPC	Agence de santé publique du Canada
CDC	<i>Centers for Disease Control and Prevention</i> (Centres de contrôle et de prévention des maladies)
CEV	Virus de l'encéphalite de Californie
CIC	Canards Illimités du Canada
COS	Classes d'occupation du sol
CPRg	Groupe <i>Culex pipiens-restuans</i>
CQP	<i>Coquillettidia perturbans</i>
CVV	Virus <i>Cache Valley</i>
EEEV	Virus de l'encéphalite équine de l'est
GC	Gouvernement du Canada
HJV	Virus <i>Highlands</i> ,
INSPQ	Institut national de santé publique du Québec
JCV	Virus <i>Jamestown Canyon</i>
LACV	Virus de l'encéphalite de La Crosse
MH	Milieux humides
MLM1	Modèle linéaire mixte à un facteur aléatoire
MLM1 <sub>sans</sub>	MLM1 sans terme d'interaction ni quadratique
MLM1 <sub>avec</sub>	MLM1 avec terme d'interaction ou quadratique
MLM1 <sub>sansx</sub>	MLM1 <sub>sans</sub> dont les variables ont été sélectionnées au seuil VIF x (x = 3, 5 ou 10)
MLM1 <sub>avecx</sub>	MLM1 <sub>avec</sub> dont les variables ont été sélectionnées au seuil VIF x (x = 3, 5 ou 10)
MSSS	Ministère de la Santé et des Services sociaux
OMS	Organisation mondiale de la santé
Pxxx	Précipitations totales moyennes du mois xxx
<i>P-P plot</i>	Diagramme probabilité-probabilité
<i>Q-Q plot</i>	Diagramme quantile-quantile
STMg	groupe <i>Ochlerotatus stimulans</i>
TNxxx	Températures minimales du mois xxx
TXxxx	Températures maximales du mois xxx
TYxxx	Températures moyennes du mois xxx
VNO	Virus du Nil occidental
SCV	Virus du sérogroupe Californie
SLEV	Virus de l'encéphalite de St-Louis
SSHV	Virus <i>Snowshoe hare</i>
VEX	<i>Aedes vexans</i>
VIF	<i>Variance Inflation Factor</i> (Facteur d'inflation de la variance)
WEEV	Virus de l'encéphalite équine de l'ouest

Un tableau récapitulatif des abréviations des noms de variables est fourni en annexe (tableau A1).



## Remerciements

Ce mémoire n'aurait sans doute pu aboutir sans la participation d'une grande équipe dont le soutien m'a été précieux tout au long de la maîtrise. Je tiens à les remercier, en commençant par Richard Fournier, mon directeur de recherche, qui m'a proposé ce sujet et qui m'a soutenue et encouragée dans les différentes étapes du projet. Je remercie également Antoinette Ludwig, ma codirectrice de mémoire, pour sa patience, sa rigueur et pour m'avoir poussée à creuser toujours plus loin la réflexion et en me servant de guide pour les étapes à suivre. Serge Olivier Kotchi, mon codirecteur, m'a permis quant à lui de survivre aux traitements titanesques des images grâce à ses compétences en programmation, ce pour quoi je lui suis infiniment reconnaissante. Pour les volets de programmation R, de statistique et de modélisation, j'ai aussi reçu l'aide salutaire de François Rousseau, de l'université de Sherbrooke. Merci également à Stéphanie Brazeau, qui travaille avec Antoinette et Serge Olivier, pour l'aide dans les démarches initiales du projet et pour son soutien moral.

Ce projet a été financé en partie grâce à l'organisme MITACS, et je remercie tous ceux qui, de près ou de loin, ont contribué à ce que mes stages soient subventionnés par le programme Acceleration. Je ne peux passer sous silence le rôle de l'équipe de GéoMont, Marylène Savoie, Jean-Philippe Sirois, Chloé Lefebvre-Dugré, qui m'ont accueillie à bras ouverts durant mes deux stages et qui m'ont tellement appris.

Je souhaite aussi souligner l'implication des membres de l'Institut national de Santé publique du Québec, dont Alejandra Irace-Cima, Marie-Ève Turcotte, Roxane Pelletier et Marie-Andrée Leblanc qui, avec Antoinette, ont travaillé fort afin que je puisse accéder aux données de captures de moustiques.

Enfin, je remercie ma famille pour le support apporté tout au long de mon retour aux études. Papa, maman, ces quelques lignes ne suffiront jamais à vous dire à quel point je vous suis reconnaissante pour tout ce que vous avez fait pour moi.

Et je ne peux clôturer cette liste sans remercier amoureusement et chaleureusement celui qui a été le phare dans mes tempêtes, le Guide et le Capitaine qui m'a aidée à garder le cap, toujours. Merci à toi Jean-Guy, pour ton amour inconditionnel, ta patience infinie et ta foi en moi, porteuse de fierté.

## Description du projet de recherche

<b>Titre du projet :</b>	Modélisation des densités moyennes annuelles de populations de moustiques associées à des maladies zoonotiques vectorielles représentant un risque pour la santé publique pour le sud du Québec
<b>Auteur :</b>	Julie Allostry
<b>Affiliation :</b>	Département de géomatique appliquée, Faculté des Lettres et des Sciences humaines, Université de Sherbrooke
<b>Programme :</b>	Maîtrise en sciences géographiques
<b>Domaine de spécialisation :</b>	Géomatique
<b>Directeur de mémoire :</b>	Richard Fournier, Université de Sherbrooke, Département de géomatique appliquée
<b>Codirecteurs :</b>	Antoinette Ludwig et Serge Olivier Kotchi, Agence de la santé publique du Canada, Laboratoire national de microbiologie
<b>Jury interne :</b>	Yannick Huot, Université de Sherbrooke, Département de géomatique appliquée
<b>Jury externe :</b>	Yves Baudouin, Université de Montréal, Département de géographie

## Avant-propos

La modélisation des densités de populations de moustiques associées à des maladies zoonotiques vectorielles représentant un risque pour la santé publique est un projet de recherche piloté par l'Agence de la santé publique du Canada, Laboratoire national de microbiologie à Saint-Hyacinthe. Le virus du Nil occidental, le virus de l'encéphalite équine de l'est ou les virus du séro groupe Californie, font partie des maladies transmises aux hommes par des moustiques qui préoccupent l'Organisation mondiale de la santé et qui sont étroitement suivis au Canada et au Québec. Modéliser puis cartographier les densités de populations des moustiques potentiellement dangereux pour l'homme est donc un enjeu de santé publique. Cela permet entre autres de guider les opérations de contrôle sur le terrain et de mieux cibler les campagnes de prévention.

## Résumé

Les changements climatiques amplifient la dispersion des moustiques et la propagation des virus dont certains moustiques sont les principaux vecteurs. Cela entraîne une augmentation du risque pour les humains d'entrer en contact avec les moustiques infectés et de développer des maladies aux conséquences parfois mortelles. Au Québec, la surveillance des variations d'abondances des espèces de moustiques vecteurs de zoonoses, comme les maladies causées par le virus du Nil occidental ou la fièvre équine de l'Est est en demande. Or, il n'existe actuellement aucun outil actif adapté au Québec permettant de prédire les abondances annuelles de populations de moustiques. Notre démarche de modélisation vise à combler ce besoin. Quatre espèces ont été étudiées dans ce projet sur la période de 2003 à 2014 pour le sud de la province : *Aedes vexans* (VEX), *Coquillettidia perturbans* (CQP) et les groupes *Culex pipiens-restuans* (CPRg) et *Ochlerotatus stimulans* (STMg). Nous avons employé une régression linéaire mixte avec un facteur aléatoire pour modéliser les abondances de chacune de ces quatre espèces de moustiques. Les variables environnementales explicatives utilisées ont été centrées réduites et incluent différentes classes d'occupation du sol ainsi que des composantes météorologiques telles que les moyennes mensuelles de précipitations totales (en millimètres) et de températures minimales, maximales et moyennes (en degrés Celsius). Par espèce, la sélection des variables explicatives statistiquement significatives à l'issue des régressions univariées s'est faite à l'aide d'une matrice de corrélation (Pearson) avec un seuil de 0,8 suivie d'un test de multicolinéarité aux seuils d'inflation de la variance, ou *Variance Inflation Factor* (VIF) de 3, 5 et 10. Puis chaque modèle correspondant à chaque seuil VIF a été testé avec et sans adjonction de termes d'interaction et de termes quadratiques. La performance de chacun des modèles a ensuite été mesurée à l'aide du critère d'information Akaike (AIC) et par l'observation du comportement des résidus (test de normalité, distance de Cook, homoscedasticité). Puis les histogrammes et cartes de dispersion des résidus de Student ont été bâtis d'une part pour les données d'entraînement, et d'autre part pour les données de validation. Nos résultats démontrent que les modèles qui prennent en compte les conditions climatiques en plus de considérer l'occupation du sol sont plus performants que les modèles qui ne considèrent que les classes d'occupation du sol, et ce, pour les quatre espèces. Pour CPRg, le meilleur modèle intègre les classes d'agriculture, de prairies et de boisés ainsi que la moyenne des températures minimales du mois de septembre de l'hiver précédent (TNsep\_1) et

le terme d'interaction Boise x TNsep\_1. Le meilleur modèle STMg est basé sur les classes urbaine et boisée, la moyenne des précipitations du mois de juin et le terme quadratique Boise x Boise. Les abondances de CQP sont modélisées à partir des milieux urbains et des précipitations moyennes des mois de janvier et d'août. Et le meilleur modèle pour VEX considère la classe d'agriculture et le terme quadratique associé Agri<sup>2</sup>, les précipitations moyennes des mois de janvier (Pjan), février et septembre (Psep) ainsi que le terme d'interaction Pjan x Psep. Les modèles se sont révélés robustes et précis sur la quasi-totalité de la zone d'étude et la présence de variables climatiques significatives pour chacune des espèces permet d'envisager la prévision à long terme des impacts des changements climatiques et de paysage sur la répartition des abondances de moustiques potentiellement nuisibles à la santé publique. En outre, c'est à notre connaissance la première étude qui modélise les abondances des espèces CQP et STMg. Notre méthode, relativement simple à reproduire, est donc applicable à différentes espèces. Nous suggérons toutefois d'utiliser des données microclimatiques et d'identifier différents paliers de densités urbaines afin de capturer plus précisément les variations environnementales locales. Il serait également intéressant de tester d'autres types de variables connues pour leur impact sur le développement ou l'activité des moustiques telles que l'humidité relative et la vitesse du vent, les données d'épandage ou la présence de prédateurs.

Mots-clés : Santé publique, moustiques, zoonoses, maladies vectorielles, modélisation spatio-temporelle, sud du Québec, *Culex pipiens-restuans*, *Ochlerotatus stimulans*, *Coquillettidia perturbans*, *Aedes vexans*.

# 1. Introduction

Au Canada, l'arbovirus le plus fréquent transmis par les moustiques est le virus du Nil occidental (VNO). Un arbovirus est un « virus habituellement transmis, dans les conditions naturelles, de vertébré à vertébré, par un arthropode hématophage, qui en constitue le vecteur » (Aubry et Gaüzère, 2017). Les moustiques font partie des arthropodes hématophages. Le premier cas humain de VNO a été identifié en 1937 en Ouganda, en Afrique de l'Est. Au Canada, les premiers cas humains ont été rapportés en Ontario puis au Québec dès 2002 sans que le virus ne soit jamais éradiqué depuis (Gouvernement du Canada (GC), 2018a). Or, bien que 70 à 80 % des personnes qui contractent le VNO restent asymptomatiques, le virus est susceptible d'entraîner des conséquences graves sur la santé des êtres humains. Du symptôme grippal à une encéphalite, voire une encéphalo-myélite sévère, le virus peut entraîner la mort dans les cas les plus graves (GC, 2018d). Les populations les plus à risque de subir des séquelles de la maladie sont celles dont le système immunitaire est faible, qui sont atteinte de maladie chronique ou encore les personnes âgées de 50 ans et plus (Ministère de la Santé et des Services sociaux (MSSS), 2018). Au Canada, deux épidémies majeures se sont déclarées en 2003 et en 2007 avec respectivement 1 481 et 2 215 cas humains recensés et en 2017, huit personnes en sont décédées. Bien que plus rares, d'autres arbovirus transmissibles aux humains sont également répertoriés sur le territoire du Canada, incluant l'encéphalite équine de l'Est (EEEV), dont le premier cas humain est apparu au Canada en 2016 (GC, 2018a), l'encéphalite de Saint-Louis (SLEV), le virus de *Cache Valley* (CVV) et les virus du séro groupe Californie (SCV). Le SCV regroupe 17 virus dont ceux de l'encéphalite de Californie (CEV), La Crosse (LACV), *Jamestown Canyon* (JCV) et *Snowshoe hare* (SSHV) (Ministère de la Santé et des Services sociaux du Québec, 2018).

L'Organisation mondiale de la santé (OMS) rapporte que les changements climatiques causent une modification dans la répartition spatiale des moustiques en favorisant leur dispersion (OMS, 2017) et leur survie dans des latitudes toujours plus septentrionales (Samy *et al.*, 2016). Une des conséquences du réchauffement climatique serait donc d'amplifier le phénomène de propagation et de développement des moustiques potentiellement vecteurs de virus pathogènes (Hoover *and* Barker, 2016; OMS, 2017; Samy *et al.*, 2016). De plus, la plupart des espèces vectrices peuvent transporter plusieurs de ces agents pathogènes à risque pour les humains (Goddard *et al.*, 2002; Harris *et al.*, 2015, Sekla *and* Stackiw, 1980). Les maladies déjà endémiques au Canada risquent

donc également de profiter de l'influence des changements climatiques en matière de prévalence ou de fréquence, ce qui pourrait augmenter les risques pour les humains.

Au Québec plus particulièrement, on recensait 20 cas humains de VNO en 2002, dont trois décès. En 2012, 134 cas ont été déclarés, dont 5 décès et en 2018, 14 décès sont survenus sur les 169 cas confirmés d'infection par le VNO. Et pour les SCV, les cas humains confirmés atteignent 100 pour 2017 et 15 en 2018. Au Québec, les SCV, l'EEEV et le VNO font partie d'une liste de vigilance destinée aux professionnels de la santé. De plus, les infections par le VNO et le SLEV sont inscrites sur la liste des maladies à déclaration obligatoire, mieux connue sous son acronyme MADO (MSSS, 2018). La recherche associée aux moustiques susceptibles de transmettre des maladies aux humains est donc un enjeu de santé publique pour le Canada en général et pour le Québec en particulier. D'où l'intérêt de développer des outils d'aide à la surveillance des populations de ces moustiques.

Les cartes de risque de densités moyennes annuelles des espèces de moustiques les plus susceptibles de transmettre les virus sous surveillance font partie d'un ensemble d'indicateurs qui contribuent à mieux cibler les épandages de larvicides (Bouden *et al.*, 2008) ou les campagnes de prévention sur le territoire. Ainsi, la cartographie par espèce constitue un outil pratique de gestion du risque pour la santé publique. La modélisation des densités de moustiques est un des moyens de cartographier ces densités de façon prédictive (Lebl *et al.*, 2013; Rochlin *et al.*, 2011; Yoo, 2014).

Un des types de modélisation les plus simples à manipuler est la régression linéaire. Ce type de modélisation a déjà été appliquée en Ontario (Mallya *et al.*, 2018) et dans les prairies canadiennes (Chen *et al.*, 2013). En modélisation liée à l'épidémiologie, diverses méthodes de régression linéaires sont employées (Cleckner *et al.*, 2011; Liu *et al.*, 2012; Ruiz *et al.*, 2010; Trawinski and Mackay, 2009; Wang *et al.*, 2011), souvent couplées à d'autres outils statistiques. Qu'il s'agisse de modèles de régression de Poisson (Brownstein *et al.*, 2002; Lebl *et al.*, 2013), de modèles logistiques multiples (Rochlin *et al.*, 2011; Schurich *et al.*, 2012), de régression multivariée *stepwise* (Liu *et al.*, 2012; Trawinski and Mackay, 2010) ou de modèles mixtes (Chen *et al.*, 2013; Yoo, 2014), c'est généralement la distribution de l'abondance des moustiques, révélée par les analyses des données de capture, qui dicte le choix de la méthode (Wang *et al.*, 2011). Les modèles linéaires nécessitent que les données respectent une distribution la plus normale possible, ce qui n'est généralement pas le cas des abondances de moustiques. Il est alors

possible d'appliquer une transformation aux données pour respecter ce critère (Chen *et al.*, 2013; Cleckner *et al.*, 2011). Plusieurs transformations doivent être testées pour chacune des espèces afin de trouver la transformation qui répondra à l'exigence de normalité. D'autres types de modèles que la régression linéaire doivent être envisagés si aucune transformation ne parvient à une distribution normale ou si les données ne sont pas suffisamment abondantes. La normalité d'un jeu de données entraîne généralement une normalité des résidus. Or, ce sont les résidus (écarts entre valeurs observées et valeurs prédites) qui doivent suivre une distribution la plus normale possible pour qu'un modèle linéaire puisse être validé. La vérification de ce postulat sera donc effectuée au cours de l'étape de validation, sans pour autant exiger que la distribution des résidus satisfasse strictement aux tests de normalité. En effet, la régression linéaire est suffisamment robuste pour qu'une légère violation de l'hypothèse de normalité soit acceptable (Zuur *et al.*, 2010).

Associée à un système d'information géographique, la modélisation est un des moyens de plus en plus utilisés pour supporter la lutte contre la propagation des arbovirus et de leurs vecteurs associés (Shuai *et al.*, 2006; Ozdenerol *et al.*, 2013). L'enjeu principal de la modélisation consiste donc, en résumé, à prévenir de futures épidémies potentiellement mortelles en produisant une cartographie des lieux à risque de développer une forte abondance de moustiques pathogènes (Clements and Pfeiffer, 2009). Pour limiter la quantité d'espèces de moustiques à étudier, une approche possible consiste à ne prendre en compte que les espèces qui présentent un potentiel vectoriel avéré et significatif pour la transmission d'au moins un des arbovirus ciblés (Pradel *et al.*, 2007). Ainsi, quatre espèces présentant un risque de transmettre des maladies aux humains ont été privilégiées en fonction de leurs potentiels vectoriels pour certains arbovirus (tableau 1), soit le groupe *Culex pipiens-restuans* (CPRg), *Coquillettidia perturbans* (CQP), le groupe *Ochlerotatus stimulans* (STMg) et *Aedes vexans* (VEX). Les espèces *Culex pipiens* et *Culex restuans*, difficiles à distinguer morphologiquement, sont très communément regroupées sous l'appellation *Culex pipiens-restuans*, bien que leurs habitats et leurs caractéristiques ne soient pas tout à fait semblables. De la même façon, STMg englobe les espèces *Ochlerotatus stimulans* et *Ochlerotatus hexodontus* dans notre jeu de données.



Tableau 1 : Arbovirus endémiques en Amérique du Nord par potentiel vectoriel des espèces de moustiques ciblées pour le projet.

Espèces		Vectrice de
CPRg	<i>Culex pipiens</i>	EEEV <sup>[4][12]</sup> , HJV <sup>[4]</sup> , LACV <sup>[8]</sup> , SLEV <sup>[7]</sup> , <b>VNO</b> <sup>[6][13]</sup>
	<i>Culex restuans</i>	EEEV <sup>[11][12]</sup> , JCV <sup>[2]</sup> , LACV <sup>[8]</sup> , SLEV <sup>[13]</sup> , <b>VNO</b> <sup>[6]</sup>
CQP		CVV <sup>[3]</sup> , EEEV <sup>[4][12]</sup> , JCV <sup>[9]</sup> , <b>VNO</b> <sup>[1]</sup> , WEEV <sup>[11]</sup>
STMg	<i>Ochlerotatus stimulans</i>	CVV <sup>[12]</sup> , JCV <sup>[2]</sup> , LACV <sup>[15]</sup> , SSHV <sup>[15]</sup>
	<i>Ochlerotatus hexodontus</i>	JCV <sup>[10]</sup> , SSHV <sup>[14]</sup>
VEX		CEV <sup>[11]</sup> , CVV <sup>[1][12]</sup> , EEEV <sup>[4][13]</sup> , HJV <sup>[4]</sup> , JCV <sup>[1][2][7][9]</sup> , LACV <sup>[5]</sup> , SLEV <sup>[7]</sup> , <b>VNO</b> <sup>[1][7][13]</sup> , WEEV <sup>[7]</sup>

CPRg : groupe *Culex pipiens-restuans* – CQP : *Coquillettidia perturbans*

STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

CVV: virus *Cache Valley* – CEV: virus de l'encéphalite de Californie – CVV: virus *Cache Valley*

EEEV: virus de l'encéphalite équine de l'est – HJV: virus *Highlands* – JCV : virus *Jamestown Canyon*

LACV: virus de l'encéphalite de La Crosse – SLEV : virus de l'encéphalite de Saint-Louis

SSHV: virus *Snowshoe hare* – VNO : virus du Nil occidental.

[1] Anderson *et al.*, 2015 – [2] Andreadis *et al.*, 2008 – [3] Andreadis *et al.*, 2014

[4] Andreadis *et al.*, 1998 – [5] Berry *et al.*, 1986 – [6] Edillo *et al.*, 2009 – [7] Goddard *et al.*, 2002

[8] Harris *et al.*, 2015 – [9] Main *et al.*, 1979 – [10] Murdock *et al.*, 2010 – [11] Sekla and Stackiw, 1980

[12] Srihongse *et al.*, 1980 – [13] Turell *et al.*, 2001 – [14] Wagner *et al.*, 1975 – [15] Walker *et al.*, 1993.

Au Québec, la principale espèce vectrice du VNO serait *Culex pipiens* (Rocheleau, 2016). Selon une étude effectuée en Illinois, les *Culex pipiens* et *restuans* affectionnent particulièrement les milieux boisés ainsi que les champs, les prairies et les zones urbaines dans une moindre mesure (Gardner *et al.*, 2014). Toujours en Illinois, Jacob *et al.* (2009 b), indiquent que les *Culex pipiens* et *Culex restuans* préfèrent les contenants d'eau artificiels et naturels que l'on trouve principalement dans les zones urbaines. VEX, une autre espèce très présente sur le territoire provincial, pond principalement dans les milieux aquatiques temporaires des zones agricoles et se développe en présence de milieux humides ou agricoles ou à proximité de milieux aquatiques (Rocheleau, 2016). D'autres sources indiquent que les VEX présentent plutôt les mares ou étangs peu profonds des pâturages, boisés et prairies pour se développer (Strickman, 1980 et 1982). Quant à elles, les larves de CQP ont besoin de plantes aquatiques telles que les roseaux ou les nénuphars dont elles puisent l'oxygène en en perceant les racines, grâce à un appendice adapté à

cette fin. C'est donc une espèce qui a besoin de tourbières ou marais permanents pour la partie larvaire de son cycle de vie (Bosak *and* Crans, 2002). Cependant, bien que les CQP soient très communs, on ne connaît pas grand-chose des autres conditions environnementales nécessaires au développement complet de cette espèce. STMg fait également partie des espèces qui semblent peu documentées au niveau entomologique alors qu'elles sont aussi vectrices de plusieurs des virus transmis par CQP, VEX ou CPRg. Il semblerait que STMg préfère les mares printanières pour se développer (Maire, 1977; Maire et Aubin, 1980; Wood *et al.*, 1979) et qu'*Ochlerotatus hexodontus* ait en plus besoin d'une température minimum plus élevée que d'autres espèces pour permettre à ses larves d'éclore, ce qui ferait d'elle une espèce plus tardive en saison (Wood *et al.*, 1979).

Toutes les espèces de moustiques n'ont ainsi pas les mêmes besoins, que ce soit en matière de milieux aquatiques permanents ou temporaires provoqués par les pluies, procurant ainsi un abri pour le stade aquatique de leur cycle de vie, de niveaux de température propres à certaines espèces ou encore de la présence d'habitats favorables au stade adulte. Cependant, il n'est pas évident de trouver des consensus sur les habitats préférentiels des différentes espèces. Il est donc indispensable que l'approche méthodologique passe par une étude des variables environnementales locales pour déterminer celles qui sont favorables à chaque espèce ciblée pour l'étude (Pritchard *et al.*, 2010), puisque la variation de densités des populations est dépendante d'une part des habitats favorables potentiels et d'autre part des conditions météorologiques. C'est pourquoi il est recommandé d'utiliser une approche qui combine au minimum les aspects spatiaux avec les aspects temporels (Lebl *et al.*, 2013; Trawinski *and* Mackay, 2009; Wang *et al.*, 2014; Yoo, 2014). Dans notre étude, les aspects spatiaux sont représentés par les classes d'occupation du sol (COS) et les aspects temporels par les conditions météorologiques et climatiques ainsi que par l'évolution du paysage avec les années. Dans la suite du texte, le terme variables environnementales englobera à la fois les COS et les conditions météorologiques.

La plupart de ces COS seront en partie alimentées par des précipitations prolongées pour former des plans d'eau temporaires qui créent des habitats nécessaires au stade larvaire des moustiques, d'où l'importance de considérer les précipitations (Gardner *et al.*, 2014), mais il s'agit aussi de situer les milieux humides (MH) permanents qui pourraient être propices à la croissance des espèces concernées, quelles que soient les conditions météorologiques. La classification strictement nécessaire pour les quatre espèces cibles pourrait se limiter à distinguer le milieu

urbain/bâti, les MH, les pâturages/prairies, et les boisés. Cependant, en l'absence d'informations exhaustives ou de réel consensus sur les habitats les plus favorables par espèce, il est prudent de considérer d'autres COS, d'autant qu'elles pourraient servir ultérieurement pour des espèces de moustiques dont nous ne traitons pas ici. En outre, les sols nus par exemple, ne semblent pas propices au développement de la plupart des espèces de moustiques. Cette information est donc précieuse à conserver pour la caractérisation des habitats propices en identifiant précisément les endroits qui ne le sont pas. On cherche donc à caractériser le territoire en COS les plus représentatives possible des habitats potentiellement favorables ou défavorables. Lorsque les classes sont déterminées, les variables nécessaires pour alimenter le modèle en sont extraites sous forme de pourcentages d'occupation du sol.

Afin de classer le territoire, l'emploi des outils de télédétection couplés à des systèmes d'information géographiques pour la cartographie des populations de moustiques offre plusieurs avantages. En effet, les territoires à l'étude dans ce type de projet de modélisation sont souvent très vastes. Effectuer des campagnes de terrain efficaces pour récolter les données de couverture du sol ou d'utilisation du sol à ces échelles est donc difficilement envisageable, ou très coûteux. Les images de télédétection offrent donc une alternative intéressante : en plus de couvrir de vastes étendues de territoires à différentes échelles, l'acquisition de ces images est rapide, plus économique que les campagnes de terrain, voire gratuite pour certains jeux de données et assure la qualité des données qui en seront extraites (Kotchi *et al.*, 2016) en plus de pouvoir être récentes, voire multitemporelles. Cette multitemporalité permet, par exemple, de suivre l'évolution d'un même paysage dans le temps, ce qui est utile dans le cas de modélisation prédictive basée sur des COS. Il existe de nombreux satellites qui offrent différentes applications en épidémiologie selon les caractéristiques de leurs capteurs respectifs dont plusieurs proposent une résolution suffisante pour appuyer une modélisation en lien avec les populations de moustiques, soit 30 m (Kalluri *et al.*, 2007). Quant aux conditions météorologiques, chaque espèce de moustiques sera plus ou moins sensible à des variations de facteurs météorologiques aussi divers que le nombre de degrés-jours, la photopériode, ou les températures intra ou interannuelles, selon ses caractéristiques biologiques et son stade de développement (Clements, 1999; Edillo *et al.*, 2009; Ozdenorol *et al.*, 2008; Ripoché *et al.*, 2019; Wang *et al.*, 2010; Wimberly *et al.*, 2014). Chaque étude utilise sa propre recette en fonction des données disponibles, du type de modèle utilisé et de la région où le modèle est appliqué. De plus, la

prudence est de mise lorsqu'il s'agit de transposer ces modèles d'une région à une autre. Par exemple, les modèles calibrés pour l'Europe ne sont pas applicables au territoire nord-américain, entre autres à cause de la différence d'architecture urbaine entre l'Europe et l'Amérique du Nord. Cette différence représente une variable qui semble avoir un impact sur les épidémies de VNO dans les milieux urbains d'Amérique du Nord et pas dans ceux d'Europe (Marcantonio *et al.*, 2015). Or, les déterminants environnementaux qui conditionnent le développement des moustiques sont nombreux et directement liés à la couverture du sol et à son utilisation ainsi qu'aux conditions météorologiques locales (Kotchi *et al.*, 2016). D'autre part, comme il est difficile de trouver un consensus sur des variables météorologiques universelles qui expliqueraient les variations d'abondance par espèce de moustiques, il est conseillé de conduire les études à l'échelle locale et au cas par cas (Marcantonio *et al.*, 2015). En outre, peu d'études travaillent sur un grand nombre d'espèces appartenant à des groupes différents. Cela complique donc le type de travail que nous nous proposons de développer dans le présent mémoire qui traite de plusieurs espèces simultanément. Devant cette difficulté et pour les raisons énoncées, notre approche est de traiter chaque espèce séparément, mais en utilisant un ensemble de variables assez vaste et diversifié pour répondre aux caractéristiques de chacune de nos quatre espèces.

## 2. Hypothèse et objectifs

Comme les conditions propices au développement des moustiques sont intimement liées à leur environnement et aux fluctuations des conditions météorologiques (Clements, 1999 ; Clements, 2000 ; Clements, 2012 ; Edillo *et al.*, 2009 ; Ozdenorol *et al.*, 2008 ; Wang *et al.*, 2010 ; Wimberly *et al.*, 2014), l'hypothèse de recherche soutenant la modélisation repose sur l'existence d'une corrélation significative entre les variations annuelles d'abondance de moustiques et les variables environnementales telles que les COS et les variables météorologiques, de façon conjointe et complémentaire.

L'objectif de ce projet est de développer une méthodologie de modélisation permettant de produire une cartographie des moyennes annuelles des densités de populations de moustiques ciblées parmi celles potentiellement menaçantes pour la santé publique. La vocation de ces cartes est de guider les différentes campagnes de prévention et de contrôle de la propagation de maladies vectorielles transmises par les moustiques pour le sud du Québec en dotant les autorités publiques d'un outil cartographique fonctionnel.

La méthode étant basée sur un modèle statistique de régression en fonction de variables météorologiques et environnementales, les quatre sous-objectifs du projet sont (1) d'établir les critères de choix des pièges à conserver pour refléter adéquatement les abondances annuelles de moustiques pour notre région d'étude ; (2) d'établir quelles sont les COS et variables météorologiques pertinentes qui caractérisent un environnement propice au développement de chaque espèce en fonction des caractéristiques entomologiques connues ; (3) de développer une modélisation prédictive des abondances annuelles basée sur ces variables météorologiques et environnementales et (4) de cartographier les densités annuelles des moustiques prédites par la modélisation. L'étude se déroule sur la période couvrant 2003 à 2016 inclusivement et couvre une partie du sud du Québec, incluant la grande région de Montréal.

## 3. Matériel

### 3.1. Échelle de travail

Un des défis de cette étude est de concilier la taille du territoire avec la taille des moustiques et leur faible distance de vol autonome. Cette distance variant d'une espèce à l'autre, Trawinski *and* Mackay (2010) et Yoo *et al.* (2016), suggèrent d'appliquer une zone tampon d'un kilomètre (1 km) autour de chaque piège à moustique sous laquelle les données environnementales sont extraites. Pour considérer les variations locales du paysage et des conditions météorologiques, des images satellitaires d'une résolution maximum de 30 m sont utilisées, tel que conseillé par Brownstein *et al.*, 2002, Clekner *et al.*, 2011, Diuk-Wasser *et al.*, 2006, Rochlin *et al.*, 2011 et Schurich *et al.*, 2014. Ceci implique de différencier les pièges dont l'emplacement varie de plus de 30 mètres d'une année à l'autre, par exemple en ajoutant un numéro séquentiel au nom des pièges.

### 3.2. Échantillonnage des moustiques

Les données de capture de moustiques ont été fournies en partie par l'Agence de la santé publique du Canada (ASPC), Laboratoire national de microbiologie à Saint-Hyacinthe et en partie par l'Institut national de Santé publique du Québec (INSPQ) selon une méthode similaire de capture. Cela représente plus de 30 espèces et plus de 4,1 millions de moustiques capturés pour un peu plus de 22 700 jours de piégeage cumulés par l'ensemble des pièges.

Deux types de pièges ont été utilisés pour les captures : des *Light Trap* (LT) au CO<sub>2</sub> et des *Gravid Trap* (GT). Les pièges de type LT attirent les femelles en quête de nourriture alors que les pièges de type GT attirent les femelles en quête d'un lieu de ponte (Trawinski *and* Mackay, 2009). Il est toutefois conseillé de n'utiliser qu'un seul type de piège selon le projet considéré, afin d'éviter les biais de mesures liées à la différence entre les méthodes de capture (Jacob *et al.*, 2009a; Karki *et al.*, 2016). En effet, le choix d'une méthode adaptée conditionne en grande partie la qualité de l'estimation de la densité des futures populations, car certains facteurs tels que le vent modifient le pouvoir d'attraction différemment selon les types de pièges. Cela signifie que les taux de capture pourraient ne pas être uniformes si plus d'un type de piège est utilisé (Karki *et al.*, 2016). Pour cette raison, il importe de considérer attentivement le type de pièges utilisé, en fonction de l'objectif (Jacob *et al.*, 2009a). De plus, les pièges à localisation fixe doivent être favorisés, afin

d'éviter que la dispersion spatiale n'engendre un biais dans l'échantillon (Clements *and* Pfeiffer, 2009; Wang *et al.*, 2014). Pour cette étude, les pièges LT, plus nombreux que ceux de type GT, ont été retenus. Le tableau 2 résume la répartition des sources de données de piégeage ainsi que leur distribution temporelle.

*Tableau 2 : Répartition temporelle des pièges en fonction de leur source et du type de piège.*

	Nb de pièges			Période de piégeage			Nb d'années de fonctionnement par piège (consécutives ou non)	Nb de jours de fonctionnement de l'ensemble des pièges
	INSPQ	ASPC	Total	INSPQ	ASPC	Total	Total	Total
<b>Light trap</b>	462	114	576	2003-2006 2013-2015	2006-2016	2003-2016	1 à 10 ans	15 603
<b>Gravid trap</b>	206	35	241	2003-2006	2006-2012	2003-2012	1 à 5 ans	7 156

ASPC : Agence de santé publique du Canada – INSPQ : Institut national de Santé publique du Québec.

À ce stade, il convient de vérifier que les genres sont séparés dans les échantillons, afin d'établir la modélisation uniquement sur la population des femelles (Cleckner *et al.*, 2011 ; Lebl *et al.*, 2013), car le comportement des moustiques femelles diffère de celui des mâles. En effet, seules les femelles cherchent à se nourrir avec le sang de leurs hôtes, dont les humains, et c'est l'attraction des hôtes mammifères que les pièges reproduisent pour piéger les femelles moustiques. En principe, les mâles ne sont pas attirés par les pièges, mais ils se retrouvent parfois dans les échantillons et c'est pourquoi il est important qu'ils soient correctement identifiés, ce qui est le cas pour nos données.

### 3.3. Variables environnementales

#### 3.3.1. Données d'occupation et utilisation du sol

Les couches de COS ont été créées à partir d'une combinaison de données satellitaires déjà classifiées et de données vectorielles. Parmi les données classifiées, des images d'Agriculture et Agroalimentaire Canada (AAC) qui couvrent la zone d'étude pour les années 2011 à 2016 inclusivement ont été utilisées. Elles sont issues d'images multi-capteurs (Landsat 5 et 8, AWiFS, DMC, SPOT, RapidEye, et RADARSAT-2 en mode *Wide*) classifiées pour l'inventaire annuel des cultures avec une exactitude globale d'au moins 85 % (AAC, 2016). Elles sont donc utiles pour délimiter les zones de prairies ou d'agriculture. Cependant, les milieux végétaux au cœur des milieux urbains sont pauvrement discriminés avec ce type de produit. Pour pallier ce défaut,

nous avons utilisé de façon complémentaire la couverture terrestre classifiée du Canada au sud de la limite forestière. Cette couverture a été bâtie à partir d'une combinaison d'images SPOT 4 et 5 acquises entre 2005 et 2010 dans quatre bandes multispectrales du visible et de l'infrarouge. Dans un souci de simplification, cette image classifiée sera désignée sous le terme SPOT dans la suite du texte. L'exactitude globale de la classification de l'image SPOT est de 71 % (Gouvernement du Canada (GC), 2019). Ces deux produits, AAC et SPOT, offrant une résolution de 30 et 20 m respectivement, celle de la classification finale correspondra à la moins précise des deux, soit 30 m. Comme les images AAC disponibles couvrent les années 2011 à 2016 inclusivement, cela induit donc un biais dans l'étude, car le paysage a évolué entre 2003 et 2011. Cependant, il a été choisi de conserver cette méthode parce qu'elle fournissait l'information la plus proche possible de la réalité tout en garantissant une uniformité de procédure de classification pour toutes les années disponibles.

Le cycle de vie des moustiques comporte un stade aquatique, ce qui se traduit par l'association de certaines espèces avec des MH (Clements, 1999 ; Edillo *et al.*, 2009 ; Verdonshot *et al.*, 2014 ; Wood *et al.*, 1979) ou avec des plans d'eau (Maire et Aubin, 1980). Or, AAC et SPOT ne sont pas classifiées pour caractériser adéquatement les surfaces aquatiques, qu'il s'agisse des MH ou des plans d'eau. C'est pourquoi il a été décidé d'ajouter des données vectorielles aux classifications issues d'images satellitaires. Les données vectorielles relatives à la caractérisation des milieux humides sont issues d'un ensemble de données dont les sources varient entre 1999 et 2005. Elles ont été acquises auprès de Canards Illimités du Canada (CIC) au 1 : 20 000 (CIC, 2018). Le Réseau hydrographique national (RHN) au 1 : 20 000, bâti à partir d'un ensemble de données dont les sources varient entre 1999 et 2014, a également été utilisé. Les données du RHN sont disponibles gratuitement sur le site GéoGratis du GC (Ressources naturelles Canada, 2017).

### 3.3.2. Données météorologiques

Puisqu'il est impossible de savoir à priori quelles variables météorologiques seront importantes pour prédire les densités annuelles de populations de moustiques, il a été décidé, dans le cadre de notre étude sur plusieurs espèces et sur un large territoire, d'utiliser des variables telles que les moyennes mensuelles et annuelles de températures et de précipitations. Cette approche a démontré un intérêt dans d'autres études (Hongoh *et al.* 2012; Johnson *et al.*, 2017). De plus, ces moyennes forment des indicateurs généralistes alors applicables à plusieurs espèces. Les relevés météorologiques quotidiens provenant des stations au sol d'Environnement Canada, interpolés



par méthode de distance inverse pondérée (Chen *et al.*, 2013; Ruiz *et al.*, 2010) et rééchantillonnées au kilomètre sur l'ensemble de la zone d'étude ont fourni les données météorologiques nécessaires pour les analyses. Une image par jour a été produite pour les années 2002 à 2015 inclusivement. Les images disponibles concernent les températures de l'air minimale, maximale et moyenne (en degrés Celsius) ainsi que les précipitations (en mm). Les relevés originaux sont accessibles et téléchargeables gratuitement sur le site du Gouvernement du Canada (GC, 2018c).

Le tableau 3 synthétise le détail des données utilisées pour en extraire les intrants de base.

*Tableau 3 : Métadonnées.*

Type de données	Source	Années d'acquisition	Précision	Intrants et unités
Données de capture de moustiques				
Décompte de captures par piège	ASPC et INSPQ	2003 à 2015	Non applicable	Nombre de moustiques par pièges agrégé par année
Données d'occupation et utilisation du sol				
Images classifiées	Inventaire annuel des cultures d'AAC	2011 à 2016 (6 images)	30 m (exactitude globale de classification de 85%)	Classes d'occupation du sol (%)  1 image par année
	Couverture terrestre SPOT classifiée du Canda au sud de la limite forestière	Acquisition entre 2005 et 2010 (1 image)	20 m (exactitude globale de classification de 71%)	
Données vectorielles	MH de CIC	Répertoriés entre 1999 et 2005 (1 shapefile)	1 : 20 000	
	RHN	Répertoriés entre 1999 et 2014 (1 shapefile)	1 : 20 000	
Données météorologiques				
Images	Relevés quotidiens de stations météorologiques au sol d'Environnement Canada	2002 à 2005	Interpolation rééchantillonnée à 1 km	Précipitations totales annuelles et mensuelles (mm) Températures minimales, moyennes et maximales annuelles et mensuelles (°C)

AAC : Agroalimentaire et Agriculture Canada – ASPC : Agence de la Santé publique du Canada

CIC : Canards Illimités du Canada – INSPQ : Institut national de Santé publique du Québec

MH : Milieux humides – RHN : Réseau hydrographique national

## 4. Méthode

### 4.1. Schéma méthodologique

La méthode suivie (figure 1) se résume en quatre grandes phases : (1) la phase exploratoire ; (2) la phase de modélisation ; (3) la phase de validation et (4) la phase cartographique.

La phase exploratoire comprend une première étape d'analyse descriptive des données de capture de moustiques et de cartographie de la répartition des pièges qui seront jugés pertinents pour, d'une part représenter les abondances annuelles de chacune des espèces et d'autre part séparer les données d'entraînement des données de validation. La seconde étape consiste à préparer les données environnementales, soit les COS et les données météorologiques, pour en extraire les variables explicatives.

La phase de modélisation sert à déterminer, parmi les variables explicatives, celles qui traduisent le mieux les variations de densités annuelles de populations de moustiques pour les données d'entraînement. Cette phase comprend cinq étapes. La première consiste à transformer les données de capture de façon à pouvoir employer une modélisation linéaire. La seconde étape concerne les régressions univariées qui permettent de comparer la performance individuelle de chacune des variables explicatives. Certaines variables sortiront significatives de cette étape, c'est-à-dire dont la *p-value* associée dans la régression linéaire est inférieure ou égale à 0,05 pour un intervalle de confiance de 95 %. L'indépendance des variables jugées significatives est testée dans la troisième étape à l'aide d'une matrice de similarité ainsi que par une analyse de multicollinéarité avec trois seuils de facteur d'inflation de la variance ou seuil VIF pour *Variance Inflation Factor*. L'étape quatre évalue le fonctionnement conjoint des variables explicatives significatives et indépendantes définies pour chacun des seuils dans une régression multivariée. Puis les éventuels termes quadratiques et/ou d'interactions sont calculés et introduits dans les modèles s'ils sont significatifs. L'étape cinq consiste à déterminer quel est le meilleur modèle en les comparant sur la base de leur performance au critère d'information Akaike (AIC) et sur la qualité de la normalité de la distribution de leurs résidus.

La phase de validation est divisée en deux étapes. D'abord celle de la validation interne durant laquelle on calcule les résidus de Student du meilleur modèle construit avec les données d'entraînement pour en étudier les histogrammes de distribution, l'homoscédasticité et mesurer la distance de Cook. Ensuite l'étape de validation externe qui consiste à utiliser les paramètres du meilleur modèle par espèce avec les variables explicatives correspondantes pour l'année de validation. C'est également à cette étape que les résidus des données de validation sont « studentisés » pour évaluer, au travers de leur distribution, la qualité de chacun des modèles.

La phase cartographique vérifie dans la première étape, pour chaque modèle, la répartition des résidus de Student dans l'espace à la recherche d'éventuels agrégats. Cette recherche se fait pour les données d'entraînement puis pour les données de validation afin de comparer les deux. Les cartes annuelles des densités de chacune des espèces dans la zone d'étude à l'aide du modèle retenu sont générées dans la seconde étape. Ces cartes prédictives annuelles représentent l'aboutissement du projet.

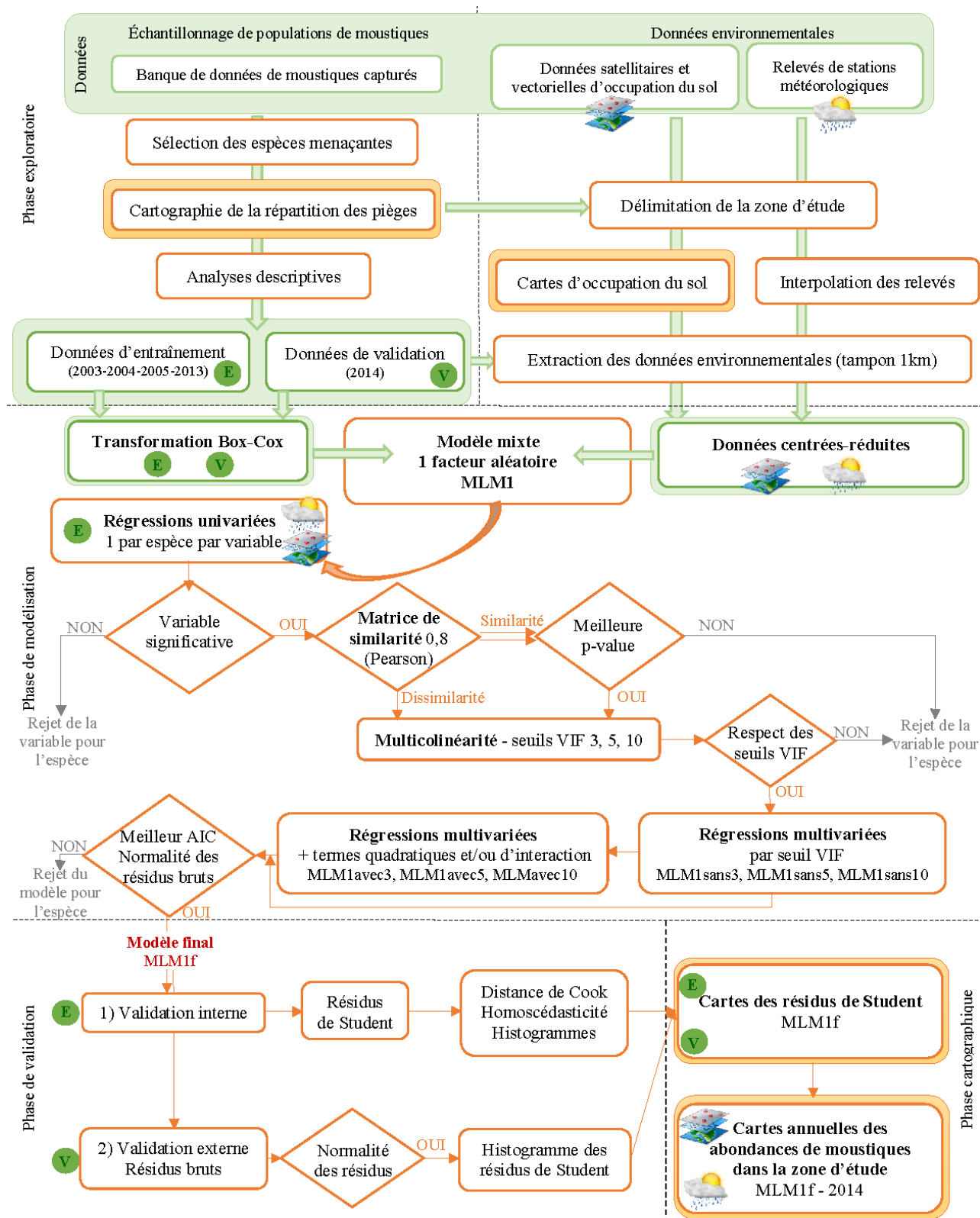


Figure 1 : Schéma méthodologique.

VIF : Facteur d'inflation de la variance – MLM1 : Modèle linéaire mixte à un facteur aléatoire –  
 MLM1avec $x$  : MLM1 avec termes quadratique et/ou d'interaction MLM1sans $x$  : MLM1 sans termes  
 quadratique ni d'interaction – MLM1f : MLM1 final.

## 4.2. Phase exploratoire

### 4.2.1. Étape 1 : cartographie de la répartition des pièges et analyses descriptives

Les analyses descriptives des données de moustiques visent à cerner les critères de choix de pièges à conserver pour refléter les densités annuelles de chacune des espèces ciblées. Comme nous avons déterminé le type de piège avec lequel travailler, nous analysons alors, pour chaque année, les histogrammes de quantités de pièges actifs en fonction de chaque semaine d'opération au cours de la saison ainsi qu'en fonction du cumul de semaines d'opération. C'est ce qui permet de déterminer le nombre de semaines minimal et suffisant de fonctionnement des pièges pour que les captures correspondantes représentent une abondance annuelle. Les semaines de capture sont caractérisées selon un système standardisé utilisé en épidémiologie et adopté à travers le monde par les Centres de contrôle et de prévention des maladies, mieux connus sous l'acronyme CDC (*Centers for Disease Control and Prevention*). Ces semaines se rencontrent sous les termes *CDC week* ou *Epiweek* (Arias, 2018).

Considérant que certaines espèces sont plus précoces ou plus tardives que d'autres en saison, nous validons la période d'activité préférentielle des espèces par les analyses de variation intra- et interannuelle de leurs densités. Ensuite, la cartographie de la répartition des pièges sur le territoire, en fonction de leur durée de fonctionnement, permet de définir l'emprise totale des pièges ayant fonctionné suffisamment longtemps. Cette emprise, à laquelle on ajoute une bordure minimalement identique aux zones tampons qui seront appliquées autour de chaque piège, soit 1 km, représente donc la zone d'étude. La cartographie, ainsi que les analyses descriptives du fonctionnement des pièges servent à distribuer les années de capture entre données d'entraînement et données de validation. Si la répartition des données l'autorise, les données de validation sont choisies parmi des années postérieures à celles des données d'entraînement (Wang *et al.*, 2011). On assure ainsi la possibilité de tester la capacité prédictive du modèle.

### 4.2.2. Étape 2 : préparation et extraction des variables environnementales

#### Préparation des classes d'occupation du sol

Parmi les COS fournies par les données retenues, nous choisissons celles qui offrent le plus d'information pertinente pour représenter les différentes COS ciblées pour le projet.

Au regard de nos objectifs de classification, nous avons bonifié chaque image de l’AAC selon les règles décrites dans le tableau 4. En effet, les milieux urbains des couches AAC ne fournissent que peu, voire aucune information sur la présence de végétation telle que les parcs et les boisés par exemple. Comme la couche SPOT est quant à elle très détaillée à ce niveau, nous récupérons les informations de SPOT en milieu urbain pour améliorer la classification d’AAC. Concernant les COS de SPOT hors milieu urbain, nous ne les considérons pas afin de que les parcelles agricoles d’AAC ne soient pas remplacées par de la simple végétation. En effet, les couches AAC permettent de distinguer clairement les types de cultures. Cependant, en dehors des pâturages et des prairies, toutes les classes distinguant les autres cultures peuvent être regroupées sous l’appellation « agriculture », car rien n’indique qu’un type de culture en particulier soit lié au développement d’une des espèces cibles. La couche des RHN a été ajoutée en dernier lieu pour identifier la classe d’eau, soient les zones d’eau profonde et peu profonde telles que les étangs, les ruisseaux et les bordures de rivières nécessaires au développement de diverses espèces. Ainsi, la classification finale retenue pour le projet comprend l’eau peu profonde et profonde, l’agriculture, les milieux humides, les prairies, les sols nus, les milieux urbains, la végétation basse et les milieux boisés, soit huit classes qui représentent autant de variables. À l’issue de cette préparation, on obtient une image bonifiée pour 2011 à 2016 inclusivement. Les zones qui ne sont pas couvertes intégralement par toutes les sources composant les COS sont exclues de la zone d’étude finale.

*Tableau 4 : Règles de classification pour bonifier la couche de base d’Agriculture et Agroalimentaire Canada (AAC).*

Si AAC ≠ urbain	Alors AAC		Bonification du milieu urbain par addition des zones végétalisées ou boisées
Si AAC = urbain et SPOT non nul	Alors SPOT	Sinon AAC	
Si AAC et SPOT et CIC = MH	Alors MH	Sinon AAC ou SPOT (classe ≠ MH)	Validation des MH
Si RHN non nul	Alors RHN	Sinon AAC	Ajout des milieux aquatiques

AAC : Agriculture et Agroalimentaire Canada – CIC : Canards Illimités du Canada – MH : Milieux humides – RHN : Réseau hydrographique national – SPOT : Couverture terrestre classifiée du Canada au sud de la limite forestière issue d’images SPOT.

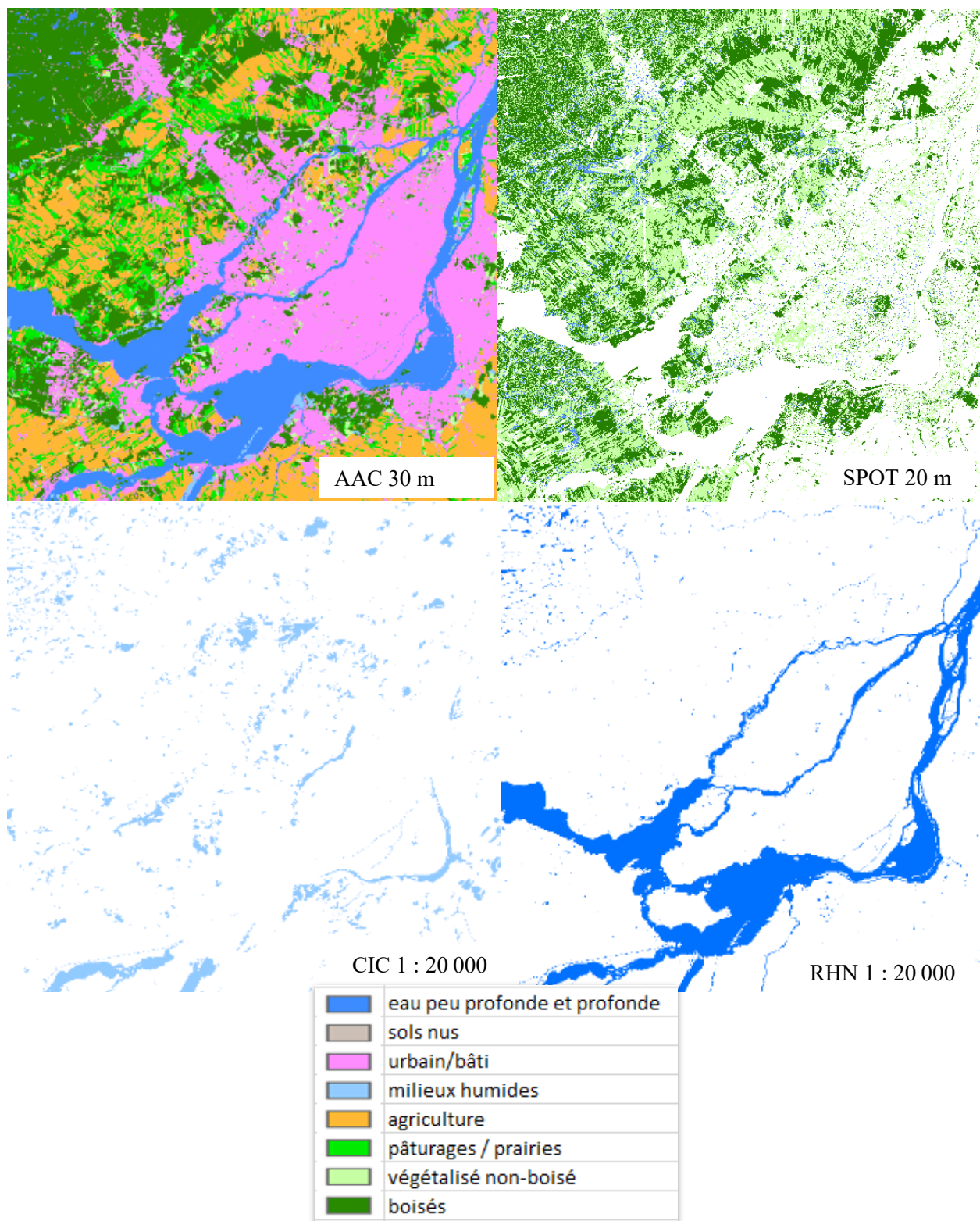


Figure 2 : Assemblage des données d'occupation du sol de base, agrandissement sur Montréal et ses alentours, 2011.

AAC : Agriculture et Agroalimentaire Canada – CIC : Canards Illimités du Canada – RHN : Réseau hydrographique national – SPOT : Couverture terrestre classifiée du Canada au sud de la limite forestière issue d'images SPOT.

Nous superposons alors les pièges sur les images classifiées pour générer les zones tampons de 1 km autour de chacun de ces pièges. Puis les pourcentages d'occupation du sol autour de chaque piège sont calculés pour chaque zone tampon. Le logiciel de traitement d'images numériques Geomatica (PCI, 2019) a été privilégié pour cette étape, car ce logiciel utilise la forme précise des polygones des zones tampons pour extraire les données. Quant à lui, le logiciel ArcGIS (ESRI, 2019) convertit les polygones en image matricielle pour effectuer les calculs. Or certains polygones de zones tampons étaient fusionnés pendant la conversion en image par manque de finesse de résolution employée par le logiciel. ArcGIS donne donc un résultat identique pour des pièges parfois éloignés de 200 m alors que Geomatica est plus précis en donnant des résultats distincts.

#### Préparation des variables météorologiques

De la même manière que pour les images de COS, on prépare les images correspondant aux variables météorologiques reconnues pour favoriser le développement des espèces de moustiques à l'étude. La figure 3 montre les précipitations moyennes de la seule journée du 7 janvier 2003. Chaque jour fournit une image de ce type, pour chacune des quatre composantes météorologiques, soit les précipitations et les températures minimales, maximales et moyennes quotidiennes. C'est à partir de ces images que l'on calcule les moyennes mensuelles et annuelles, sur lesquelles on superpose les zones tampons de chacun des pièges précédemment sélectionnés. C'est ainsi que l'on extrait la valeur moyenne de chacune des variables qui entrera dans le processus de modélisation.

Remarque : les limites administratives de la figure 3 ont été ajoutées dans le seul but d'aider le lecteur à situer spatialement les variations de précipitations moyennes quotidiennes sur le territoire. Elles n'ont aucune utilité dans la préparation des variables météorologiques.



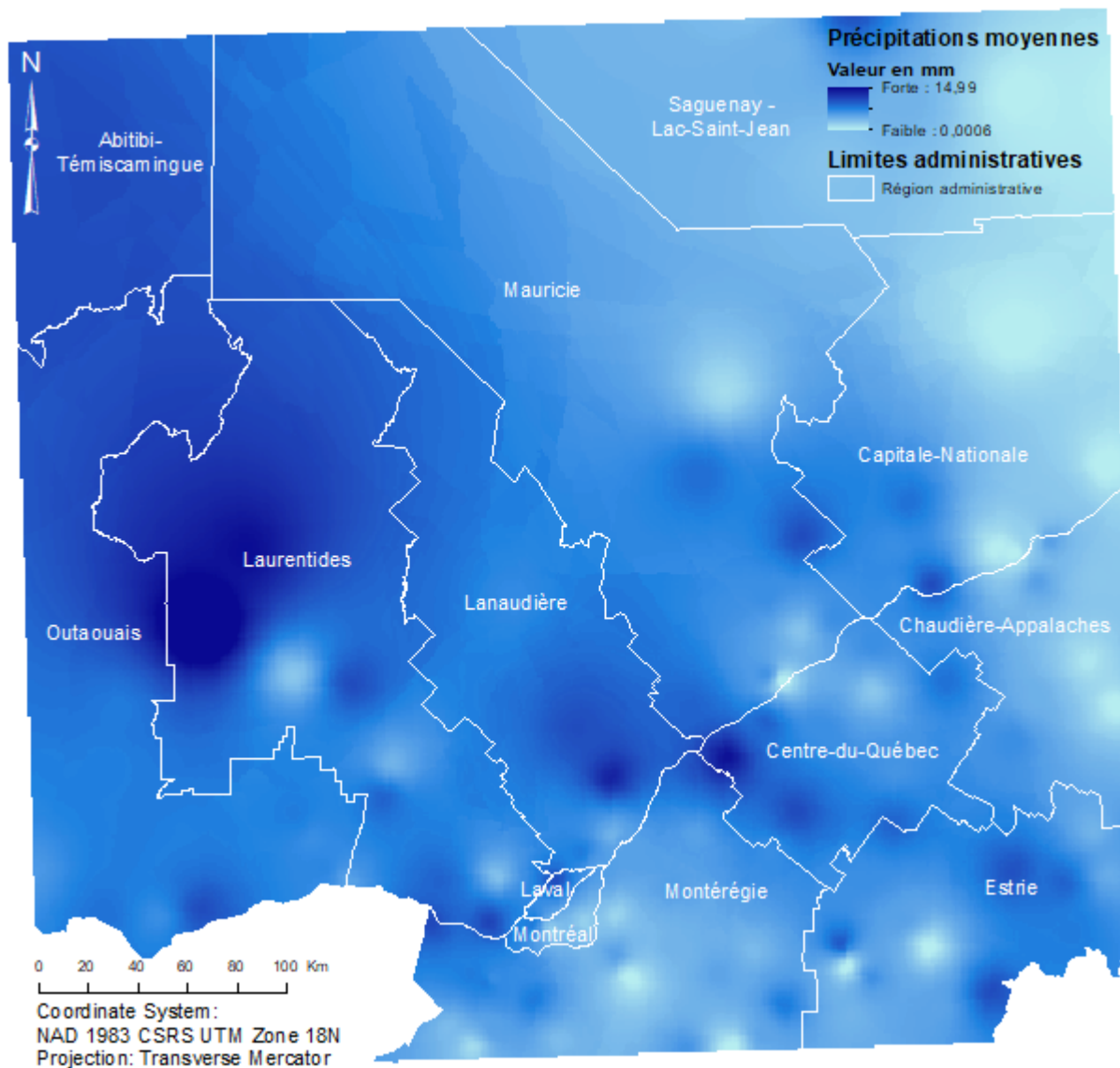


Figure 3 : Image des précipitations moyennes du 7 janvier 2003 avec contour des régions administratives.

### 4.3. Phase de modélisation

La majorité des analyses statistiques a été effectuée à l'aide du module XLSTAT (Addinsoft, 2019) pour Excel. Les résultats intermédiaires ont été validés en comparant les résultats de R avec ceux de XLSTAT afin d'en garantir l'exactitude.

Nous utilisons une approche de régression statistique de type modèle mixte (linéaire) à un facteur aléatoire (MLM1) pour modéliser la densité annuelle de chacune des espèces ciblées. L'introduction d'un facteur aléatoire dans une régression linéaire permet de prendre en compte le fait que l'échantillonnage n'est pas exhaustif. En effet, les données de piégeage ne représentent

pas l'ensemble du territoire, il ne s'agit que d'un échantillon de données possibles. Chaque piège est déjà identifié par un nom, généralement sous forme de code de type AAA 000, qui lui est propre. Cependant, sur le terrain, les pièges ne sont pas toujours replacés exactement au même endroit d'une année à l'autre. Nous ajoutons donc un numéro séquentiel à tous les pièges. Par défaut, le numéro séquentiel est zéro (AAA 000\_0). Si la localisation du piège diffère de plus de 30 m une autre année, le numéro est porté à 1 et ainsi de suite et, à l'inverse, si un piège donné a les mêmes coordonnées plusieurs années de suite, son numéro séquentiel ne change pas. Le nom du piège est ainsi composé de son nom de base et du numéro séquentiel que nous lui attribuons. C'est sur ce nom du piège séquentiel que le facteur aléatoire est porté.

#### 4.3.1. Étape 1 : transformation des données

##### Données de moustiques

L'utilisation d'un MLM1 demande une distribution normale des données de moustiques. À ce stade, nous testons différentes transformations afin de répondre à cette exigence. Parmi les transformations possibles, nous avons testé sans succès la racine carrée, l'inverse ( $10^{-1}$ ), des transformations logarithmiques ( $\log_{10}$ ,  $\ln$ ) et finalement la transformation Box Cox qui a été retenue. Chaque transformation Box Cox reçoit un coefficient de transformation Lambda ( $\lambda$ ) qui est l'exposant auquel élever les données afin que la série suive une courbe normale. Ce paramètre est calculé par le logiciel pour être optimal en fonction de la série d'origine.

##### Données environnementales

Afin de les rendre comparables lors de leur introduction dans un MLM1 multiple, les variables extraites des COS et des images météorologiques sont centrées réduites. Cela signifie que, pour chaque élément de la série de valeurs appartenant à une variable explicative donnée, on soustrait l'espérance puis on divise par l'écart-type de cette série de valeurs.

#### 4.3.2. Étape 2 : régressions univariées

Pour chaque espèce, nous évaluons l'association linéaire de chaque variable environnementale au sein des zones tampons avec les abondances de moustiques correspondantes. Dans un intervalle de confiance de 95 %, si la *p-value* de la variable testée est inférieure à 0,05, on considère que l'association est statistiquement significative pour l'espèce. Elle est donc conservée pour les prochaines étapes de la modélisation. Les variables non significatives sont rejetées.

### 4.3.3. Étape 3 : indépendance des variables

Toutes les variables significatives ne peuvent être introduites dans le MLM1. Il faut s'assurer au préalable que certaines d'entre elles ne sont pas dépendantes les unes des autres. Parmi les méthodes de recherche de corrélation entre différentes variables, on retrouve la corrélation de Pearson, ou matrice de similarité-dissimilarité, qui permet de réduire le nombre de variables en entrée du MLM1 sans que la qualité n'en soit compromise (Chen *et al.*, 2013; Rochlin *et al.*, 2011; Schurich *et al.*, 2012). Le seuil conseillé varie entre 0,75 (Rochlin *et al.*, 2011; Schurich *et al.*, 2012) et 0,8 (Chen *et al.*, 2013) que nous avons retenu. Ainsi, si deux variables sont similaires à 80 % ou plus, elles sont comparées entre elles et celles dont la *p-value* est la plus faible sont conservées.

À l'issue de ce premier tri, les variables restantes subissent un test de multicolinéarité. Il s'agit d'éliminer les variables qui ont une relation linéaire entre elles. Pour cela, on utilise le VIF (Cleckner *et al.*, 2011; Zuur *et al.*, 2010). Il est employé afin d'éviter une redondance dans les régressions linéaires multivariées. Au-delà d'un certain seuil, la variable est corrélée aux autres variables selon une combinaison linéaire (Foucart, 2006). On considère alors que la participation de cette variable dans le MLM1 deviendrait négligeable comparativement aux autres et on l'élimine.

Un seuil VIF inférieur à 10 est acceptable pour éliminer les colinéarités significatives (Brunkard *et al.*, 2007; Dohoo *et al.*, 2003). On trouve cependant des seuils VIF d'une valeur de 5 pour ce type de modélisation (Cleckner *et al.*, 2011), voire encore plus restrictifs avec une valeur fixée à 3 (Zuur *et al.*, 2010). Nous choisissons d'appliquer ces trois seuils VIF, ce qui générera théoriquement trois modèles par espèce. Cela permet d'évaluer la performance de chacun des modèles résultants et de les comparer pour ne conserver que le modèle le plus efficace.

### 4.3.4. Étape 3 : régressions multivariées

Les variables sélectionnées aux seuils de VIF 3, 5 et 10 sont introduites dans le MLM1. Pour cela nous appliquons la méthode de l'algorithme descendant pour en comparer les résultats avec ceux de la méthode *step-wise* ou pas-à-pas (Liu and Weng, 2012). Cette seconde méthode est conseillée pour éliminer un éventuel problème de multicolinéarité résiduelle entre les variables (Foucart, 2006). On s'assure ainsi que les modèles soient robustes au sens statistique.

Pour l'algorithme descendant, nous introduisons toutes les variables explicatives retenues dans le MLM1 et nous observons leur comportement lorsqu'elles travaillent ensemble. Si certaines variables deviennent non-significatives, nous comparons leur coefficient de participation dans l'explication des variations de densités de moustiques, et nous éliminons celle dont le coefficient est le plus faible. Puis nous relançons le MLM1 sans cette variable explicative. Le processus est réitéré jusqu'à ce que toutes les variables explicatives du MLM1 soient significatives, soit avec une *p-value* inférieure ou égale à 0,05.

Nous confrontons ensuite le résultat final avec les résultats issus d'un processus *step-wise* ascendant. Cette approche commence avec une seule variable, soit celle ayant la meilleure *p-value* et le plus fort taux de participation dans la régression univariée. Nous introduisons ensuite une seule nouvelle variable, et cela pour chaque variable explicative significative issue des analyses de multicollinéarité. Il y a donc autant de régression multivariée à deux variables que de variables explicatives, moins celle du départ. Le MLM1 conservé est celui dont la seconde variable obtient la meilleure *p-value* et/ou le coefficient de participation le plus fort si la *p-value* est identique. Puis nous introduisons une troisième variable selon le même procédé et ainsi de suite jusqu'à ce qu'il ne reste plus aucune variable explicative à introduire ou jusqu'à ce qu'une variable devienne non significative dans le MLM1. Un résultat identique selon les deux processus renforce l'assurance de la robustesse du modèle final. Nous effectuons ce processus pour chacune des espèces et pour chacun des groupes de variables explicatives déterminés par chacun des trois seuils VIF, soit théoriquement trois modèles par espèce. À ce stade, les trois MLM1 par espèce sont encore au stade intermédiaire (MLM1sans<sub>3</sub>, MLM1sans<sub>5</sub> ou MLM1sans<sub>10</sub>).

Pour chacune des variables explicatives composant les MLM1sans, on effectue une recherche de tendance qui puisse être expliquée par un terme quadratique. En effet, pour la plupart de ces variables, la relation univariée avec la densité de moustiques n'est pas purement linéaire. C'est-à-dire que le nuage de points représentant le nombre de moustiques par piège en fonction de la variation d'une variable en particulier ne suit pas une droite parfaite, mais plutôt une tendance linéaire qui, dans certains cas, peut être décrite plus précisément en introduisant un terme quadratique ou un terme d'interaction. Un terme quadratique se traduit par le carré de la variable. Les termes quadratiques éventuels sont donc calculés puis centrés réduits, de la même façon que les variables explicatives pures. Ces termes quadratiques doivent cependant être significatifs pour entrer dans la modélisation. Ainsi, chacun des termes quadratiques calculés subit une régression

linéaire univariée à 1 facteur aléatoire, en fonction des quantités de moustiques de l'espèce correspondante. Nous considérons donc le terme quadratique pour l'introduire dans le MLM1sans correspondant si et seulement si sa *p-value* est significative au terme de la régression univariée.

De la même façon, les termes d'interaction sont calculés et centrés réduits pour chacune des espèces. Un terme d'interaction est le résultat d'une multiplication entre deux variables explicatives composant le MLM1sans. Nous croisons ainsi entre-elles toutes les variables explicatives qui participent à chacun des MLM1sans obtenus pour calculer tous les termes d'interaction possible au sein de chaque MLM1sans. Comme pour les termes quadratiques, nous ne considérons la variable pour l'introduire dans le MLM1sans que si sa *p-value* est inférieure ou égale à 0,05 au terme d'une régression univariée pour l'espèce correspondante.

Nous relançons alors la régression avec ces éventuels termes quadratiques et/ou d'interaction significatifs comme variables explicatives en plus des variables explicatives d'origine du MLM1sans correspondant. Les termes quadratiques et/ou d'interaction ne sont conservés que si ces derniers ainsi que la ou les variables explicatives d'origine dont ils sont issus restent significatifs.

Nous obtenons donc potentiellement six modèles par espèce ne comportant que des variables explicatives significatives et indépendantes pour l'espèce : le modèle de base intermédiaire MLM1sans dont les variables explicatives sont sélectionnées selon chacun des trois seuils VIF (MLM1sans<sub>3</sub>, MLM1sans<sub>5</sub>, et MLM1sans<sub>10</sub>) plus éventuellement trois modèles agrémentés de termes quadratiques et/ou d'interaction (MLM1avec<sub>3</sub>, MLM1avec<sub>5</sub> et MLM1avec<sub>10</sub>). Chaque modèle est en fait une formule mathématique du type  $y = ax + bz (+...) + c$ , où  $y$  représente le nombre de moustiques par piège (le cas échéant, en valeur transformée) et comportant autant d'arguments centrés réduits ( $x$ ,  $z$ ,...) que de variables explicatives environnementales, plus une constante appelée *intersect* ( $c$ ). Chacune des variables possède un coefficient de participation au modèle ( $a$ ,  $b$ ,...), plus ou moins élevé en fonction de l'importance de l'influence de la variable pour prédire les abondances de moustiques. Comme cette influence contribue positivement ou négativement, le coefficient de participation est, selon le cas, soit positif, soit négatif.

#### 4.3.5. Étape 5 : choix du meilleur modèle

Pour mesurer la performance de chacun des modèles obtenus, nous employons l'*Akaike Information Criterion* (AIC) (Chen *et al.*, 2013; El Adlouni *et al.*, 2007; Rochlin *et al.*, 2011; Schurich *et al.*, 2014). L'AIC définit comme meilleur modèle celui qui possède la log-vraisemblance prédite maximisée, c'est-à-dire celui dont la distance Kullback-Leibler entre une distribution ajustée et vraie est minimisée (Sakamoto, 2018). Chacun des modèles testés reçoit un score numérique correspondant. On cherche donc le score minimum pour choisir le meilleur modèle par espèce.

Parallèlement, le comportement des résidus de chaque modèle est vérifié statistiquement et graphiquement pour en valider la distribution normale (Wajnberg, 2011; Zuur *et al.*, 2010). L'évaluation statistique est réalisée à l'aide des tests de Shapiro-Wilk, Anderson-Darlin, Lilliefors et Jarque-Bera, tandis que l'évaluation graphique sommaire se fait par le biais des diagrammes probabilité-probabilité (*P-P plot*) et quantile-quantile (*Q-Q plot*) dans lesquels les résidus doivent être le plus alignés possible avec la première bissectrice du graphique. Le meilleur modèle, ou modèle final (MLM1f), est celui dont la distribution des résidus est la plus proche de la normale et qui obtient le meilleur score AIC.

### 4.4. Phase de validation

#### 4.4.1. Étape 1 : validation interne

Nous testons maintenant la robustesse du modèle sélectionné. L'homoscédasticité est évaluée visuellement et sommairement à l'aide d'un graphique de répartition. Pour construire le nuage de points, on oppose les valeurs prédites (en abscisse) aux résidus de Student (en ordonnée). Comme les résidus ne doivent dépendre ni des variables explicatives ni des valeurs d'observation, leur dispersion doit donc être homogène et ne pas présenter de structure particulière. Si la distribution des résidus d'un modèle s'éloigne trop de la normalité et que leur dispersion démontre de l'hétéroscédasticité, nous vérifions si des données d'observation extrêmes, voire aberrantes, ne viennent pas biaiser l'estimation des coefficients de participation des modèles.

Les valeurs extrêmes sont également étudiées au travers des histogrammes des résidus de Student dont les valeurs doivent majoritairement rester dans l'intervalle  $[-2 ; 2]$ . Une quantité importante de valeurs en dehors de l'intervalle démontrerait que le modèle ne reproduit pas correctement

certaines des variations dans les abondances de populations de moustiques. Une proportion de 95 % des résidus dans cet intervalle est acceptable.

Afin de vérifier la présence de valeurs extrêmes, nous calculons dans R la distance de Cook des résidus de Student, qui mesure l'impact de chaque observation dans la régression. La présence d'extrêmes est considérée comme problématique si la valeur de la distance de Cook associée à ce résidu dépasse le seuil critique de 1. Le cas échéant, le graphique résultant met en lumière les valeurs résiduelles dont l'observation correspondante altère potentiellement la qualité du modèle. Ce pourrait être le cas si un piège a mal fonctionné pendant une certaine période, capturant ainsi moins de moustiques que prévu, ou dans le cas d'une erreur humaine de retranscription ou de décompte dans les données d'observation par exemple. Il est alors possible d'enlever ces données d'observation et de relancer l'étape de modélisation sans elles pour évaluer son impact réel sur l'estimation des coefficients de participation des variables explicatives. Cependant, il faut prendre en compte qu'en régression multivariée, juger une valeur comme étant aberrante est plus délicat que dans les cas de régression univariée (Cleroux *et al.*, 1990). De plus, l'étude des variations d'abondance des populations est un phénomène naturel. Il est alors admis que quelques extrêmes font partie de la norme et il n'est donc pas conseillé d'enlever systématiquement les observations jugées à priori aberrantes (Zuur *et al.*, 2010). La distance de Cook nous permet néanmoins de valider la robustesse d'un modèle, tout comme l'analyse des histogrammes et de la distribution des résidus de Student.

#### 4.4.2. Étape 2 : validation externe

L'étude des résidus des années d'entraînement étant satisfaisante, on calcule maintenant les abondances transformées de moustiques avec les variables explicatives de validation. Les valeurs prédites sont comparées aux valeurs observées et on analyse alors les résidus des données de validation en testant la normalité de leur distribution. Comme pour les données d'entraînement, on calcule ensuite les résidus de Student pour construire leurs histogrammes. Ainsi, plus le modèle est précis, plus les valeurs standardisées resteront comprises dans l'intervalle  $[-2; 2]$  et plus la moyenne sera proche de zéro. Ici encore, si un modèle surestime ou sous-estime les abondances de moustiques, cela n'affecte pas sa qualité. Cependant, l'information recueillie sur le comportement des résidus et leur comparaison entre espèces pourrait se révéler intéressante s'il existe des similarités.

## 4.5. Phase cartographique

### 4.5.1. Étape 1 : carte des résidus des modèles

Comme il existe quelques valeurs de résidus en dehors de l'intervalle critique, on cartographie leur répartition spatiale à la recherche d'éventuels groupements ou d'une structure particulière. Le cas échéant, que ce soit pour les données de validation ou les données d'entraînement, cela soulève des questions. En effet, si un modèle présente une répartition aléatoire des résidus extrêmes, alors on peut supposer que les variations d'abondance sont globalement bien estimées. Alors que des groupements ou une structure tendraient à démontrer que le modèle est passé à côté d'un phénomène explicatif de la variabilité des abondances de moustiques de l'espèce concernée dans la région concernée.

### 4.5.2. Étape 2 : cartes annuelles des densités d'espèces dans la zone d'étude

Si toutes les étapes de validation sont jugées satisfaisantes, il est possible de cartographier les abondances annuelles de moustiques à l'échelle de la zone d'étude. Chaque carte met alors en lumière les sites les plus à risque de voir éclore une abondante population de moustiques présentant un risque pour la santé publique. Pour les construire, nous reproduisons les conditions dans lesquelles les modèles ont été créés, à savoir, centrer-réduire les valeurs des images représentant les variables explicatives de COS et de météo de chaque modèle.

Pour un modèle ayant une COS comme variable explicative, on utilise l'image correspondant à l'année dont on souhaite obtenir les prédictions d'abondances des espèces de moustiques. Dans notre étude, l'image 2011 étant la plus récente disponible, les données qui en sont extraites fournissent les variables pour les années 2003 à 2005 et nous utilisons les images 2013 et 2014 pour produire les variables des années correspondantes. Cependant, chacune de ces images est binaire, elle ne contient que deux valeurs : 0 % ou 100 %. En effet, soit un pixel représente la COS en question et porte la valeur 100, soit le pixel ne correspond pas à la COS et sa valeur est nulle. Par exemple, pour la classe « urbain », dans toute la zone d'étude, chaque pixel de l'image est soit urbain, soit il appartient à une autre classe. Dans le processus, les pourcentages de COS proviennent d'une zone tampon d'1 km autour de chacun des pièges. En imaginant qu'un pixel donné représente l'emplacement d'un piège, on doit donc attribuer à ce pixel le pourcentage de COS-variable explicative dans un rayon d'1 km, ce qui revient à effectuer une moyenne. En appliquant ce principe à toute l'image, nous créons un effet de bordure de 500 m autour de la



zone d'étude dans laquelle les valeurs de pixels sont faussées. Nous excluons donc cette bordure pour la suite. L'image couvrant la zone d'étude rognée comporte maintenant des valeurs potentiellement incluses entre 0 et 100 %, en supposant qu'un pixel pourrait être isolé de la COS sur 1 km à la ronde (0 %) ou au contraire en plein cœur d'une zone couverte par la COS en question sur 1 km à la ronde (100 %). Ensuite, toutes les valeurs doivent être centrées réduites, selon les mêmes paramètres que ceux employés lors de la première étape de la phase de modélisation (ces transformations sont illustrées par les figures 39 à 41 du paragraphe 5.4.2). C'est-à-dire qu'on utilise les valeurs de moyenne et d'écart-type obtenus avec les données d'entraînements lors de la phase de modélisation et non ceux qui seraient calculés avec les données environnementales de l'année de validation.

Concernant les images météorologiques, les données ont été interpolées à partir de celles fournies par les stations au sol. La précision des images est donc proportionnelle à la proximité entre deux stations. Comme la distance entre deux stations atteint parfois plusieurs dizaines de kilomètres, la précision de l'image est inférieure au kilomètre qui représente la taille des pixels. De plus, les valeurs de ces pixels sont continues au fil de l'image. Il est donc inutile de moyenner chacun des pixels des images météo sur une distance de 1 km avant de les centrer-réduire, contrairement aux données d'occupation du sol dont la précision est de 30 m et dont les valeurs sont discrètes.

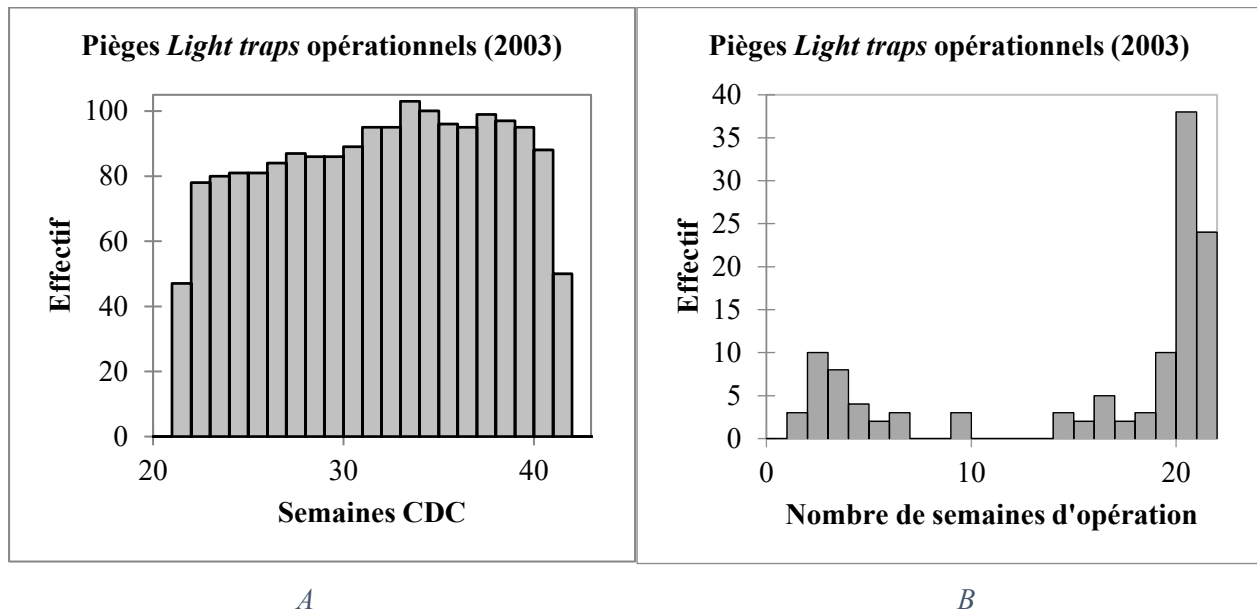
Nous reprenons alors la formule mathématique de type  $y = ax + bz (+ \dots) + c$  de chaque MLM1f, soit une formule par espèce que l'on applique aux images centrées réduites des variables explicatives correspondantes (les images correspondent aux  $x, z, \dots$  ; l'*intersect* et les coefficients de participation sont les paramètres fixes du modèle). C'est le résultat de ce calcul d'images qui forme la cartographie annuelle des abondances transformées Box Cox de populations de moustiques par espèce.

## 5. Résultats

### 5.1. Phase exploratoire

#### 5.1.1. Étape 1 : cartographie de la répartition des pièges et analyses descriptives

Les analyses descriptives des données de moustiques amènent à la conclusion que seuls les pièges de type *Light Trap*, les plus nombreux, ayant fonctionné au minimum 15 semaines consécutives sont conservés. Les semaines sont traditionnellement décomptées selon le système CDC. En 2003, la semaine 21 commençait le 18 mai et la semaine 41 commençait le 11 octobre (figure 4A). Une période plus longue est trop restrictive dans le nombre de pièges candidats et compromettrait la suite des analyses. Une période plus courte n'est pas suffisante pour représenter les abondances annuelles. Certaines années ne répondant pas à l'exigence de 15 semaines consécutives, elles ont été écartées de l'étude. Pour l'exemple, le fonctionnement des pièges de l'année 2003, décrit par les histogrammes 4A et 4B, démontre qu'au plus 102 pièges offrent des données de capture. Ces pièges ont fonctionné de façon globale entre la 21<sup>e</sup> et la 41<sup>e</sup> semaine CDC inclusivement (figure 4A), dont 84 seulement ont fonctionné un minimum de 15 semaines (figure 4B).



Figures 4A et B : Histogramme de la période d'opération des pièges de type *Light trap* de 2003 (A) et histogramme de la durée d'opération des pièges de type *Light trap* de 2003 (B).

Les années qui comportent suffisamment de pièges répondant aux critères de sélection sont 2003, 2004, 2005, 2013 et 2014, qui ont été utilisées comme données d'entraînement. L'année 2014 a donc été sélectionnée pour fournir les données de validation. Cela correspond à 184 entrées par espèces pour 123 pièges distincts pour les données d'entraînement et 206 entrées par espèce par piège pour 2014. Bien que l'écart entre les années d'entraînement et de validation soit relativement faible pour rechercher un éventuel impact des changements climatiques, 95% des données d'entraînement proviennent des années 2003 à 2005 et l'évolution du paysage est remarquable. Ainsi, cette répartition entre entraînement et validation a été motivée dans le but de tester la valeur prédictive des modèles, en réservant une année postérieure aux années d'entraînement pour la validation, en espérant que les résultats puissent être utilisés dans un contexte où des données météorologiques seraient issues des modèles de simulation des changements climatiques.

La cartographie de la répartition des pièges en fonction de leur durée de fonctionnement détermine la zone d'étude finale (figure 5). Les limites de cette zone s'étendent des latitudes 45°00 N à 48°74 N et des longitudes -76°44 W à -70°64 W. Pour permettre les calculs effectués lors d'analyses spatiales, toutes les données ont été projetées dans le système Universal Transverse Mercator, North American Datum 1983 zone 18 nord (UTM NAD83 18N). La figure 6A montre la répartition des pièges en fonction de leur durée de fonctionnement pour 2004 et la figure 6B celle de 2014. Les cartes des autres années d'entraînement et des années rejetées sont présentées en annexe.

La zone d'étude finale aurait pu être tronquée à la hauteur de la ville de Québec, puisqu'il n'y a pas de données de piégeage plus au nord (figure 5). Cependant, avant d'opter pour la méthodologie décrite dans ce mémoire, une approche à l'échelle hebdomadaire avait été testée et toutes les images avaient été préparées pour prendre en compte les données de pièges posés au Saguenay que l'on aperçoit dans la figure 6A. Afin de ne pas perdre de temps à traiter de nouveau toutes les images déjà prêtes pour l'extraction des variables, il a été décidé de conserver l'étendue telle qu'illustrée par la figure 5.

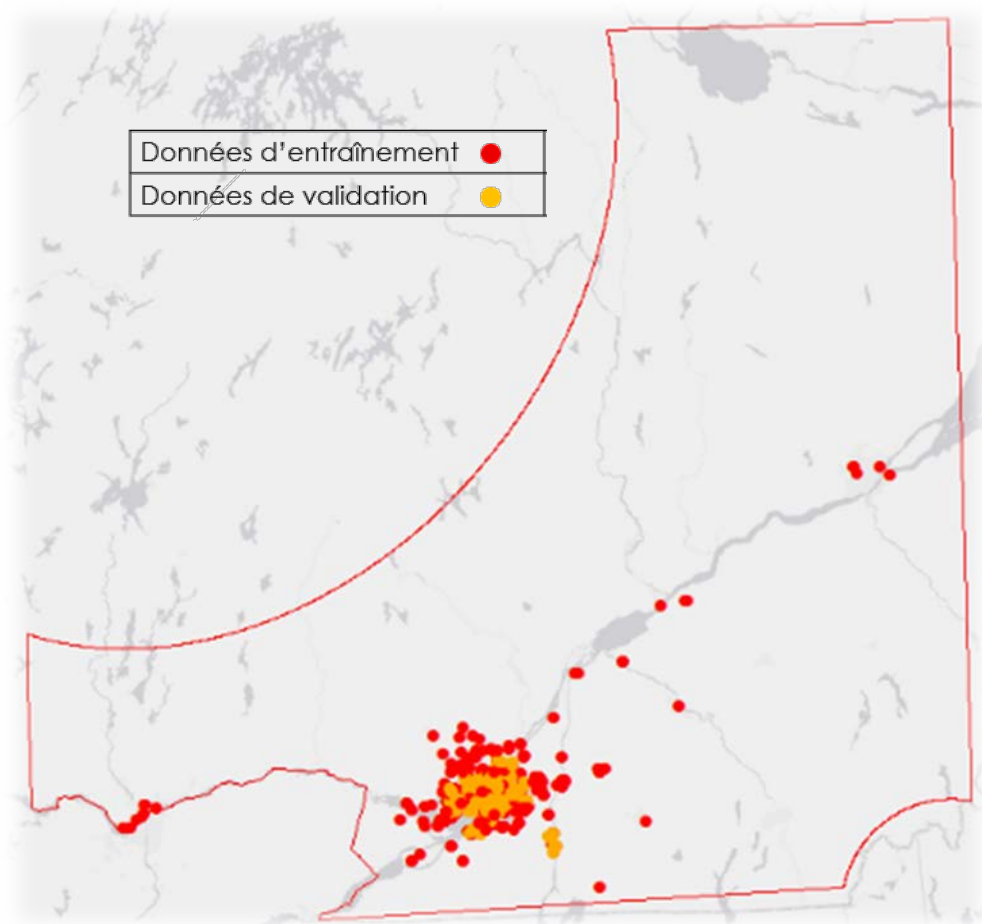
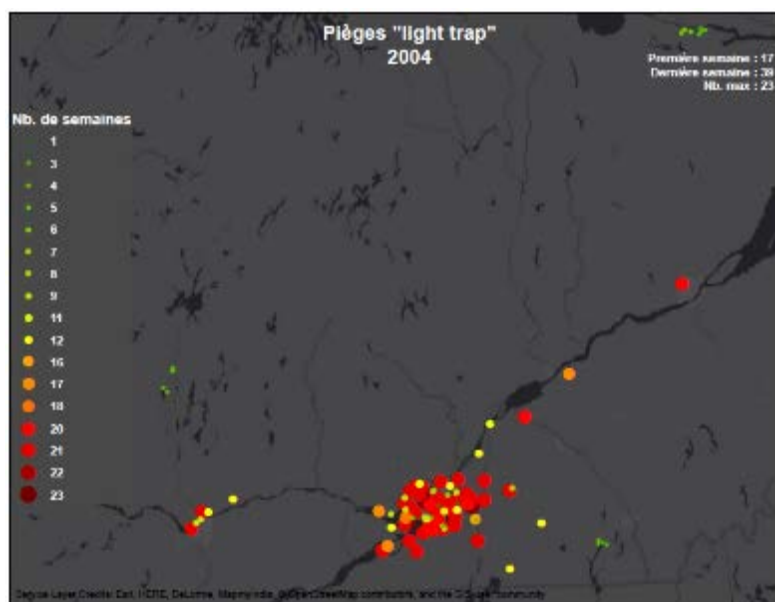
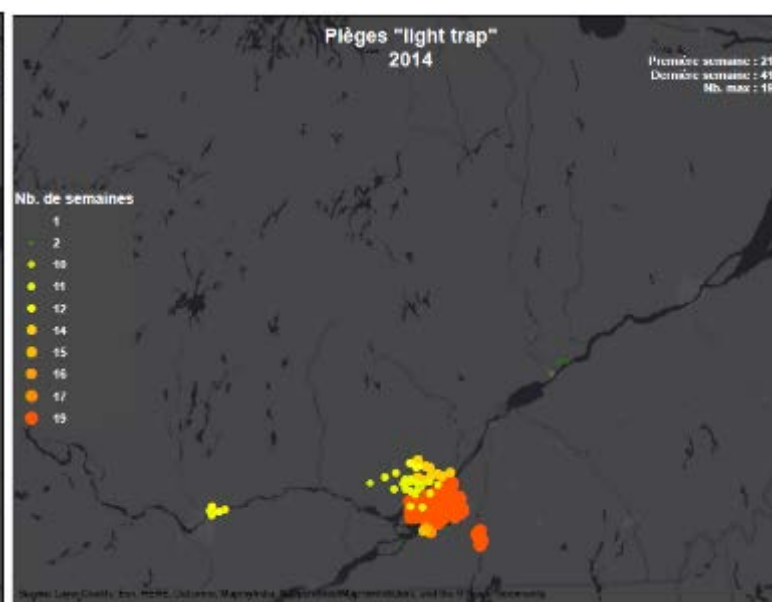


Figure 5 : Délimitation de la zone d'étude.



6A



6B

Figures 6A et B : Répartition des pièges selon leur durée de fonctionnement pour 2004 et 2014.

## 5.1.2. Étape 2 : variables environnementales

### 5.1.2.1. Variables d'occupation du sol

Notre bonification des couches AAC apporte une amélioration de la qualité de classification visuellement appréciable dans les milieux urbains (figure 7), là où se situe la majorité des pièges. Le nombre de pixels occupés par chacune des huit COS sous la zone tampon de 1 km autour de chacun de ces pièges est retranscrit sous forme de tableaux. C'est le résultat de la conversion de ces nombres qui forme la banque de variables explicatives (tableau 5).

*Tableau 5 : Extrait des pourcentages d'occupation du sol sous les zones tampons de 1 km autour des pièges.*

Site_seq	Annee	pct11_1km_eau	pct11_1km_sol	pct11_1km_urbain	pct11_1km_Mhtot	pct11_1km_agri	pct11_1km_prairie	pct11_1km_vege	pct11_1km_boise
DOL 104_0	2011	0,00	0,00	87,42	0,00	0,23	0,97	8,85	2,52
DOL 105_0	2011	0,00	0,00	89,35	0,00	0,23	0,83	8,76	0,83
DOR 001_1	2011	0,00	1,52	73,22	0,00	0,00	0,00	24,63	0,63
DOR 001_2	2011	0,00	1,55	72,60	0,00	0,00	0,00	25,25	0,60
DOR 101_0	2011	0,00	1,43	74,15	0,00	0,00	0,23	22,58	1,60
DOR 102_0	2011	0,03	0,00	73,60	0,00	0,00	0,52	21,09	4,76
DOR 103_0	2011	0,03	0,00	68,86	0,00	0,00	0,86	26,56	3,69
DRU 001_0	2011	20,52	0,17	47,91	0,17	0,17	0,23	0,63	30,19
EST 100_0	2011	0,00	0,46	73,63	0,00	2,21	2,43	14,52	6,76
EST 101_0	2011	0,00	0,69	68,15	0,00	3,67	2,09	14,72	10,68
EST 102_0	2011	26,94	0,00	56,09	0,46	0,00	0,11	10,63	5,76

agri : classe agriculture – Mhtot : classe regroupant l'ensemble des milieux humides – pct : pourcentage  
Site\_seq : nom et numéro séquentiel du piège.

Bien que les milieux humides soient faiblement représentés, toutes les classes sont couvertes, plus ou moins selon la localisation des pièges. La classe majoritaire est sans équivoque ni surprise le milieu urbain. Dans l'échantillon de résultats du tableau 5, on constate par exemple que l'environnement des pièges varie du urbain-eau (EST 102\_0) au urbain-végétalisé non boisé (DOR 001\_2) en passant par le urbain-boisé-eau (DRU 001\_0). Ce sont les éventuelles relations entre cette variabilité au sein de l'environnement proximal des pièges et la variation des abondances de moustiques capturés par chacun de ces pièges que nous tentons de découvrir.

### 5.1.2.2. Variables météorologiques

97 % des pièges ont fonctionné au minimum jusqu'à la dernière semaine de septembre incluse. Les moyennes pour l'année en cours s'étalent donc de janvier à septembre de l'année en cours et de septembre à décembre de l'année précédente, soit 13 images par type de relevé météo (4) pour chaque année d'entraînement plus quatre images (4) pour les moyennes annuelles par type de relevé météo. Nous obtenons donc 56 [(13 x 4) + 4] images météorologiques différentes par année, auxquelles on superpose les mêmes zones tampons que pour les COS afin d'en extraire les variables correspondant à chacun des pièges sous forme de tableau.



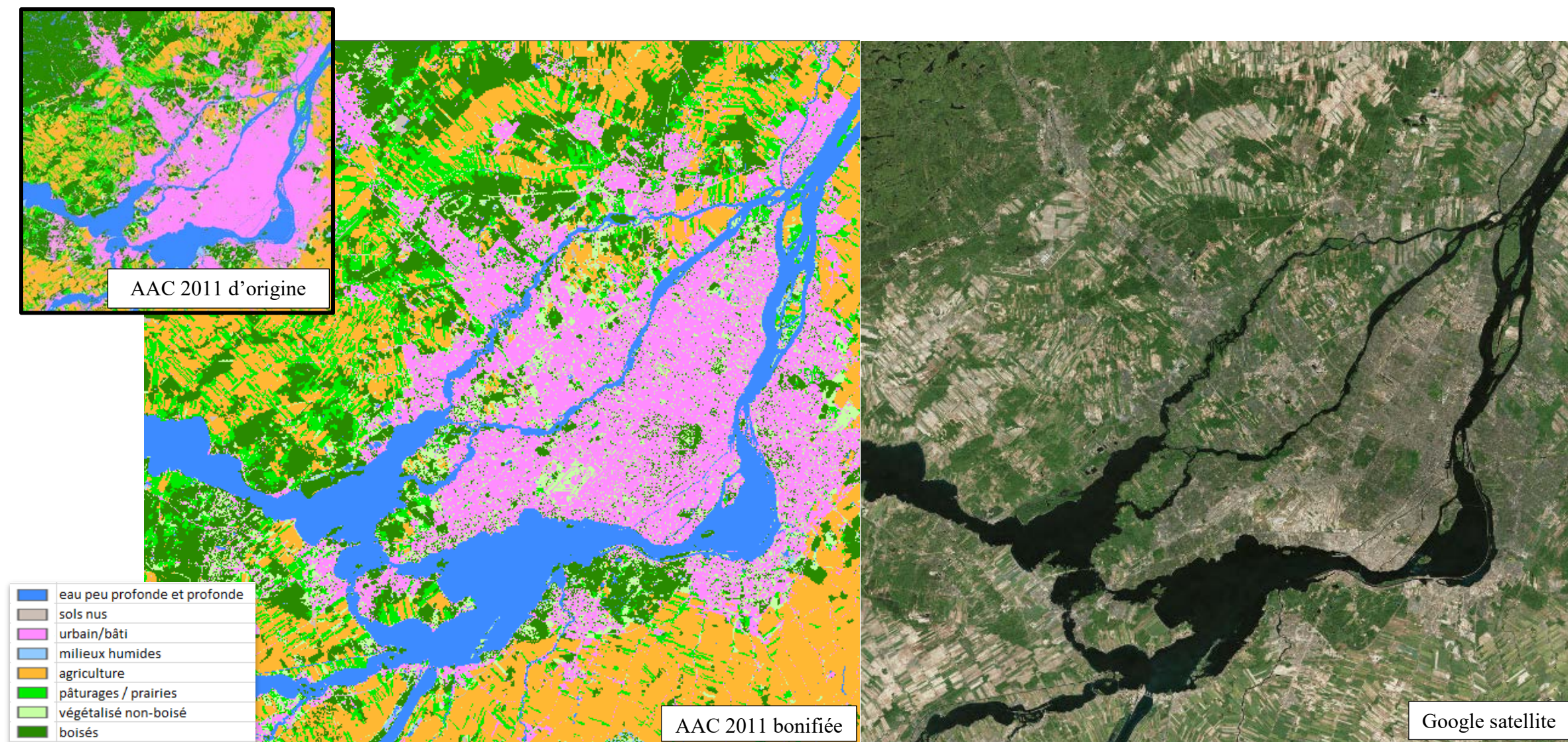


Figure 7 : Couche des classes d'occupation du sol bonifiée de 2011 – comparaison avec une image satellite de Google Earth – agrandissement sur l'île de Montréal.  
AAC : Agriculture et agroalimentaire Canada.



## 5.2. Phase de modélisation

### 5.2.1. Étape 1 : transformation des données

#### Données de capture

Pour les quatre espèces, c'est la transformation Box Cox qui a permis de rapprocher les distributions d'abondance des données de capture annualisées d'une structure normale. Comme la transformation Box Cox n'accepte pas les valeurs nulles, il a fallu les convertir en une valeur fictive très faible ( $10^{-6}$ ) qui différencie clairement les quantités nulles de la quantité 1. Cela touche 2 % des données de STMg et aucune des autres espèces. Les coefficients lambda calculés par XLSTAT pour chaque espèce ainsi que le décompte des valeurs d'abondance nulles sont récapitulés dans le tableau 6.

Les abondances annuelles transformées Box Cox pour chacun des pièges satisfont majoritairement aux tests de normalité pour les quatre espèces, sauf STMg qui échoue au test Shapiro-Wilk (tableau 7). Comme ce sont surtout les comportements des résidus qui importent, la quasi-normalité de la distribution des STMg Box Cox nous permet malgré tout l'emploi d'une modélisation de type régression linéaire, à laquelle le nom des pièges est ajouté comme facteur aléatoire (MLM1). À titre comparatif, les tableaux 8 et 9 présentent les statistiques descriptives avant et après transformation Box Cox des données de capture par espèce.

Tableau 6 : Paramètres de la transformation Box Cox des données de capture.

Nb de valeurs nulles				
	CPRg	CQP	STMg	VEX
Entraînement	0	0	7/184	0
Validation	0	0	1/207	0
Lambda ( $\lambda$ ) de la transformation Box Cox				
	CPRg	CQP	STMg	VEX
Entraînement	0,069	-0,099	0,155	-0,124

CPRg : groupe *Culex-piapiens restuans* – CQP : *Coquillettidia perturbans*

STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

Tableau 7 : Test de normalité des données d'entraînement transformées Box Cox.

Variable\Test	Shapiro-Wilk	Anderson-Darling	Lilliefors	Jarque-Bera
CQP	0,547	0,440	0,237	0,435
VEX	0,126	0,370	0,664	0,327
CPRg	0,674	0,677	0,459	0,907
STMg	<b>0,012</b>	0,065	0,451	0,257

CPRg : groupe *Culex-piapiens restuans* – CQP : *Coquillettidia perturbans*

STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

Tableau 8 : Statistiques descriptives des données d'entraînement d'origine.

Statistique	CPRg	CQP	STMg	VEX
Nb. d'observations	184	184	184	184
Minimum	3	2	0	45
Maximum	2315	41 923	11 215	17 425
Effectif du minimum	1	2	7	1
Effectif du maximum	1	1	1	1
Amplitude	2312	41 921	11 215	17 380
1er Quartile	81,50	32,00	20,00	377,75
Médiane	192,00	103,00	85,00	941,50
3e Quartile	410,25	400,75	373,00	2646,00
Somme	61 035	182 590	72 948	447 888
Moyenne	331,71	992,34	396,46	2434,17
Variance (n-1)	151 574,14	16 357 955,78	971 464,24	13 494 870,56
Écart-type (n-1)	389,36	4044,50	985,63	3673,54

CPRg : groupe *Culex-piapiens restuans* – CQP : *Coquillettidia perturbans*

STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

Tableau 9 : Statistiques descriptives des données d'entraînement transformées Box Cox.

Statistique	CPRg	CQP	STMg	VEX
Nb. d'observations	184	184	184	184
Minimum	1 142	0,670	-5 703	3 036
Maximum	10 256	6 587	20 885	5 671
Effectif du minimum	1	2	7	1
Effectif du maximum	1	1	1	1
Amplitude	9 114	5 918	26 587	2 634
1er Quartile	5 145	2 935	3 811	4 205
Médiane	6 343	3 719	6 387	4 620
3e Quartile	7 466	4 524	9 691	5 036
Somme	1 167 608	680 952	1 202 799	849 670
Moyenne	6 346	3 701	6 537	4 618
Variance (n-1)	2 712	1 448	22 046	0,305
Écart-type (n-1)	1 647	1 203	4 695	0,552

CPRg : groupe *Culex-piapiens restuans* – CQP : *Coquillettidia perturbans*

STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

### Variables environnementales

Nous avons centré réduit les huit variables de COS et les 56 variables météorologiques. Parmi ces variables environnementales, les tableaux 38 et 39 du paragraphe 5.5.2 présentent les moyennes et écart-types des variables significatives sélectionnées par les modèles finaux.

#### 5.2.2. Étape 2 : régressions univariées

Le sommaire des *p-values* de chacune des 64 variables, pour chacune des espèces est présenté en annexe. Analysée individuellement à l'aide de la régression linéaire et dans un intervalle de confiance de 95 %, la répartition des variables environnementales significatives par type et par espèces est récapitulée dans le tableau 10.



Tableau 10 : Nombre de variables significatives par espèce au terme des régressions univariées.

Nb de variables significatives	CPRg	CQP	STMg	VEX
COS	3	3	3	2
Météorologiques	21	36	35	43
<b>Total</b>	<b>24</b>	<b>39</b>	<b>38</b>	<b>45</b>

CPRg : groupe *Culex-pipiens restuans* – CQP : *Coquillettidia perturbans*

STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

### 5.2.3. Étape 3 : indépendance des variables

En ne conservant que la *p-value* la plus significative entre deux variables similaires à 80 % et plus pour l'espèce en question, il reste 15 variables non corrélées pour CPRg, 22 pour CQP, 20 pour STMg et 25 pour VEX. On constate que toutes les variables de COS issues des régressions univariées sont indépendantes les unes des autres. Aucune n'a été éliminée entre les résultats récapitulés dans le tableau 10 et ceux présentés dans le tableau 11.

À ce stade, seul VEX comporte plus de variables liées aux précipitations que de variables reliées aux températures. Cependant, les disparités entre nombre de variables de températures et nombre de variables de précipitations sont faibles comparativement à la quantité moindre de COS.

Au terme du test de multicollinéarité menée avec des seuils VIF fixés à 3, 5 et 10, il ne subsiste aucune variable météorologique sous le seuil VIF 3 (tableau 12), pour aucune des quatre espèces. Or, un modèle sans variable météorologique ne peut atteindre un objectif de prédiction lié aux changements climatiques. Toutefois, dans le but de comparer le niveau de performance entre un modèle avec et sans variables météorologiques, la modélisation a été malgré tout conservée avec un seuil de VIF 3.

Tableau 11 : Variables centrées réduites significatives par espèce au terme du test de similarité à 80 %.

CPRg	CQP	STMg	VEX
Agri	Urbain	Urbain	Eau
Prairie	Agri	Prairie	Agri
Boise	Boise	Boise	TXan
TXan	TXan	TNjan	TNavr
TNsep_1	TNavr	TNfev	TNmai
TXfev	TNmai	TNmar	TNjun
TXjun	TNaou	TNavr	TNdec_1
TXaou	TNdec_1	TNmai	TXfev
TXoct_1	TXfev	TNjun	TXaou
TXnov_1	TXjun	TNjul	TXsep
TYavr	Txsep_1	TNoct_1	TXsep_1
Pmar	TXoct_1	Txsep	TXoct_1
Pavr	TXnov_1	Txsep_1	TXnov_1
Pmai	TYjul	Pjan	Pjan
Paou	Pjan	Pmar	Pfev
	Pmar	Pjun	Pmar
	Pavr	Paou	Pavr
	Pmai	Psep	Pmai
	Pjul	Pnov_1	Pjun
	Paou	Pdec_1	Pjul
	Psep		Paou
	Pdec_1		Psep
			Psep_1
			Poct_1
			Pdec_1

CPRg : groupe *Culex pipiens restuans* – CQP : *Coquillettidia perturbans*  
STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

Tableau 12 : Variables centrées réduites explicatives significatives et indépendantes par espèce au terme du test de multicolinéarité avec le sens de corrélation, positif ou négatif.

CPRg	CQP	STMg	VEX
+ Agri	+ Agri	– Urbain	– Eau
– Prairie	– Urbain	+ Prairie	+ Agri
– Boise	+ Boise	+ Boise	– Pfev
– TNsep_1	+ Pjan	– Pjan	– Psep
– Pmar	+ Psep	+ Pmar	+ Pjan
+ Paou	+ Pjul	– Pjun	+ Pjul
	+ Paou	– Paou	+ Psep
		– Psep	

VIF <=10  
VIF <=5  
VIF <=3

CPRg : groupe *Culex pipiens restuans* – CQP : *Coquillettidia perturbans*  
STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

Pour CPRg, aucune variable n'est conservée avec une valeur de VIF entre 3 et 5, et c'est la seule espèce pour laquelle une variable liée aux températures reste significative. Il s'agit de la moyenne des températures minimales du mois de septembre de l'année précédente (TNsept-1) qui est corrélée négativement avec les quantités de moustiques. Toutes les autres variables météorologiques sont exclusivement des données de précipitation. Concernant les COS, il n'y a que CQP pour laquelle deux variables, urbain et boisé, se situent dans un seuil supérieur à 3. En considérant le seuil VIF 10, toutes les espèces sont dotées de trois variables reliées aux COS, sauf VEX qui n'en compte que deux.

CPRg compte seulement deux variables dont l'influence est positive, soit l'agriculture et les précipitations du mois d'août, contrairement aux quatre autres variables qui sont corrélées négativement. Pour CQP, toutes les variables ont une influence positive sur son abondance sauf l'urbain qui est la seule qui semble être défavorable à son abondance et qui influence également négativement les STMg. STMg comporte trois variables corrélées positivement contre cinq qui défavorisent son abondance alors que VEX compte trois variables corrélées négativement contre quatre positivement.

Les variables météorologiques associées aux densités de moustiques sont très différentes d'une espèce à l'autre. En considérant que l'hiver se déroule entre décembre et janvier, CQP s'avère sensible aux conditions hivernales (précipitations de janvier), estivales (précipitations de juillet et août) et automnales (précipitations de septembre), mais pas aux conditions printanières. Le constat est identique pour VEX avec les précipitations de janvier et février en hiver, les précipitations du mois de juillet pour l'été et les précipitations du mois de septembre de l'année en cours et de l'année précédente pour l'automne. Les abondances de CPRg ne semblent pas affectées par les conditions hivernales, mais plutôt par celles du printemps (précipitations de mars) en plus de celles de l'été (précipitations du mois de mars) et de l'automne (TNsept-1). Quant à STMg, toutes les saisons semblent importantes avec les précipitations de janvier en hiver, celles de mars au printemps, de juin et août pour l'été et de septembre pour l'automne. Il est curieux de constater qu'en hiver, ce sont les précipitations de janvier qui paraissent prédominer, et celles d'août et septembre pour l'automne.

Pour les COS, l'agriculture semble prédominante puisqu'elle est associée positivement de façon statistique à trois des quatre espèces (CQP, VEX et CPRg), puis l'urbain qui est corrélé négativement pour CQP et STMg, les prairies et les boisés corrélés tous deux négativement pour

CPRg et positivement pour STMg, et enfin l'eau, corrélée négativement pour VEX uniquement. Aucune des autres COS ne se montre influente sur leur abondance.

#### 5.2.4. Étape 4 : régressions multivariées

Un nouveau palier de sélection des variables s'effectue en appliquant chacun des trois seuils VIF. Ainsi, les régressions multivariées, effectuées par le biais de l'algorithme descendant, génèrent trois modèles de base par espèce, soit un par seuil VIF. La seule exception est pour CPRg qui ne compte que deux modèles puisqu'aucune de ses variables significatives ne correspond au seuil de VIF 5. Cela représente donc 11 modèles intermédiaires dont les paramètres sont récapitulés dans les tableaux 13 à 23. Afin de distinguer les modèles selon les différents seuils, ils seront appelés MLM1sans3, MLM1sans5 et MLM1sans10 dans la suite du texte. Le terme « sans » indique que le modèle ne comporte ni terme quadratique, ni terme d'interaction.

- Groupe *Culex pipiens-restuans*

*Tableau 13 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance (MLM1sans3) pour le groupe Culex pipiens-restuans.*

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,376	0,118	59	53,839	< 0,0001	6,142	6,610
Agri	0,721	0,120	59	5,990	< 0,0001	0,483	0,958
Prairie	-0,394	0,123	59	-3,217	0,002	-0,636	-0,152
Boise	-0,324	0,123	59	-2,626	0,011	-0,568	-0,081

*Tableau 14 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 10 de variation d'inflation de la variance (MLM1sans10) pour le groupe Culex pipiens-restuans.*

#	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,369	0,118	58	53,961	< 0,0001	6,136	6,602
Agri	0,677	0,122	58	5,567	< 0,0001	0,437	0,916
Prairie	-0,420	0,123	58	-3,430	0,001	-0,662	-0,179
Boise	-0,340	0,123	58	-2,763	0,008	-0,584	-0,097
TNsep 1	-0,219	0,105	58	-2,094	0,041	-0,426	-0,013

Dans ces deux modèles intermédiaires de CPRg, toutes les variables de COS sont restées significatives. En revanche, les variables de précipitations ont toutes été éliminées du

MLM1sans10. Il ne subsiste qu'une seule variable de température, soit la température minimum du mois de septembre de l'année précédente (TNsep\_1). En dehors de l'agriculture dont la *p-value* est constante à <0,0001, celle de la prairie baisse très légèrement de 0,002 dans le MLM1sans3 à 0,001 pour le MLM1sans10. La *p-value* du boisé décroît de 0,003 du premier modèle au second. Quant à la valeur absolue des coefficients de participation de prairie et boisé, celle-ci augmente. Cela semble démontrer que l'introduction d'autres variables dans le modèle renforce le rôle des variables de COS dans la prédiction des abondances de CPRg pour au moins deux d'entre-elles.

- *Coquillettidia perturbans*

Tableau 15 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance (MLM1sans3) pour *Coquillettidia perturbans*.

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	3,703	0,102	61	36,291	< 0,0001	3,502	3,905
Agri	0,229	0,098	61	2 349	0,022	0,037	0,422

Tableau 16 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 5 de variation d'inflation de la variance (MLM1sans5) pour *Coquillettidia perturbans*.

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	3,714	0,088	59	42,143	< 0,0001	3,540	3,887
Urbain	-0,508	0,087	59	-5 844	< 0,0001	-0,680	-0,337
Pjan	0,293	0,065	59	4 528	< 0,0001	0,165	0,420
Psep	0,193	0,070	59	2 760	0,008	0,055	0,331

Tableau 17 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 10 de variation d'inflation de la variance (MLM1sans10) pour *Coquillettidia perturbans*.

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	3,700	0,091	59	40,693	< 0,0001	3,521	3,879
Urbain	-0,535	0,090	59	-5 947	< 0,0001	-0,713	-0,358
Pjan	0,244	0,062	59	3 904	0,000	0,121	0,367
Paou	0,265	0,070	59	3 769	0,000	0,126	0,403

Pour CQP, il n'y avait qu'une seule variable significative au seuil VIF 3, l'agriculture. Le MLM1 est donc logiquement le résultat de la régression univariée de la variable agriculture. Nous constatons que pour les MLM1sans5 et MLM1sans10, l'agriculture a été éliminée au profit de l'urbain qui est significative dans les deux modèles ainsi que les précipitations du mois de janvier. La seule distinction se situe dans la dernière variable que sont les précipitations du mois de septembre pour MLM1sans5 contre le mois d'août pour MLM1sans10, dont le coefficient de participation est plus élevé que pour les précipitations du mois de janvier, ce qui n'est pas le cas pour le MLM1sans5.

- Groupe *Ochlerotatus stimulans*

*Tableau 18 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance (MLM1sans3) pour le groupe Ochlerotatus stimulans.*

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,532	0,348	60	18,766	< 0,0001	5,845	7,219
Urbain	-1,588	0,490	60	-3,240	0,002	-2,555	-0,621
Boise	1,090	0,496	60	2,197	0,032	0,111	2 069

*Tableau 19 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 5 de variation d'inflation de la variance (MLM1sans5) pour le groupe Ochlerotatus stimulans.*

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,490	0,342	59	18,999	< 0,0001	5,816	7,164
Urbain	-1,589	0,481	59	-3,307	0,002	-2 538	-0,641
Boise	1,079	0,487	59	2,217	0,030	0,119	2,039
Pjan	-0,691	0,262	59	-2 639	0,011	-1 208	-0,175

*Tableau 20 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 10 de variation d'inflation de la variance (MLM1sans10) pour le groupe Ochlerotatus stimulans.*

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,518	0,336	59	19,377	< 0,0001	5,855	7,182
Urbain	-1,500	0,473	59	-3,169	0,002	-2,434	-0,566
Boise	1,056	0,479	59	2,203	0,032	0,110	2,001
Pjun	-1,160	0,241	59	-4,805	< 0,0001	-1,636	-0,684

Pour STMg, chacun des trois modèles sélectionne l'urbain et le boisé comme variables de COS significatives. La prairie n'est donc pas restée significative dans les MLM1sans5 et 10. La seule différence qui persiste au terme de cette étape est la variable météorologique représentée par les précipitations du mois de janvier pour MLM1sans5 et celles du mois de juin pour MLM1sans10. Ici aussi, le coefficient de participation des précipitations moyennes est plus élevé dans le MLM1sans10 que dans le MLM1sans5. La *p-value* des variables de COS reste stable entre les trois modèles.

- *Aedes vexans*

*Tableau 21 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance (MLM1sans3) pour Aedes vexans.*

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	4,621	0,041	60	111,479	< 0,0001	4,539	4,702
Eau	-0,106	0,042	60	-2,540	0,014	-0,188	-0,024
Agri	0,176	0,041	60	4,319	< 0,0001	0,096	0,256

*Tableau 22 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 3 de variation d'inflation de la variance (MLM1sans5) pour Aedes vexans.*

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	4,619	0,039	59	119,136	< 0,0001	4,542	4,695
Eau	-0,105	0,039	59	-2,687	0,009	-0,182	-0,028
Agri	0,166	0,039	59	4,306	< 0,0001	0,090	0,242
Pfev	-0,130	0,036	59	-3,588	0,001	-0,202	-0,059

*Tableau 23 : Modèle linéaire mixte à une variable aléatoire dont les variables significatives simples ont été sélectionnées au seuil 10 de variation d'inflation de la variance (MLM1sans10) pour Aedes vexans.*

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	4,622	0,036	57	127,609	< 0,0001	4,551	4,694
Eau	-0,084	0,036	57	-2,290	0,026	-0,155	-0,012
Agri	0,140	0,036	57	3,923	0,000	0,070	0,211
Pfev	-0,172	0,037	57	-4,717	< 0,0001	-0,244	-0,100
Pjan	0,169	0,035	57	4,813	< 0,0001	0,100	0,238
Psep	0,128	0,034	57	3,811	0,000	0,062	0,195

Les trois MLM1sans de VEX conservent les deux variables de COS comme significatives bien que leur coefficient de participation décroisse entre les seuils 3 et 10. Le comportement du MLM1sans10 diffère de celui des autres espèces dans le sens où toutes les variables du MLM1sans3 restent significatives dans le MLM1sans5, dont on retrouve aussi toutes les variables dans le MLM1sans10. Quant au coefficient de participation des précipitations de février, il augmente entre le MLM1sans5 et le MLM1sans10. La *p-value* de l'eau semble être la plus affectée par l'introduction d'autres variables comparativement aux autres *p-value* qui oscillent entre 0 et  $<0,0001$ .

En suivant la démarche *step-wise* ascendant, on retrouve le MLM1sans10 pour chacune des quatre espèces. Cela renforce notre degré de confiance en la robustesse de la méthode puisque le résultat est identique et laisse également supposer que les MLM1sans10 devraient être plus performants que les MLM1sans3 ou MLM1sans5. Nous choisissons de poursuivre malgré tout les analyses avec les 3 seuils VIF pour le valider.

Les régressions univariées des variables composant les MLM1sans montrent pour la plupart une tendance dans les nuages de points mettant en relation les quantités de moustiques transformées Box Cox en abscisse et les variables explicatives centrées réduites en ordonnée (figures 8 à 11). Bien que la répartition des nuages de points n'en démontre pas clairement l'influence, les termes quadratiques retenus sont  $\text{agri}^2$  pour CPRg et VEX,  $\text{boise}^2$  pour STMg et aucun pour CQP. Pour les espèces STMg et VEX, l'influence du terme quadratique est négative. Elle est positive pour CPRg. Les termes quadratiques retenus sont donc calculés ainsi que tous les termes d'interaction possibles, puis ils sont centrés réduits afin d'être testés.



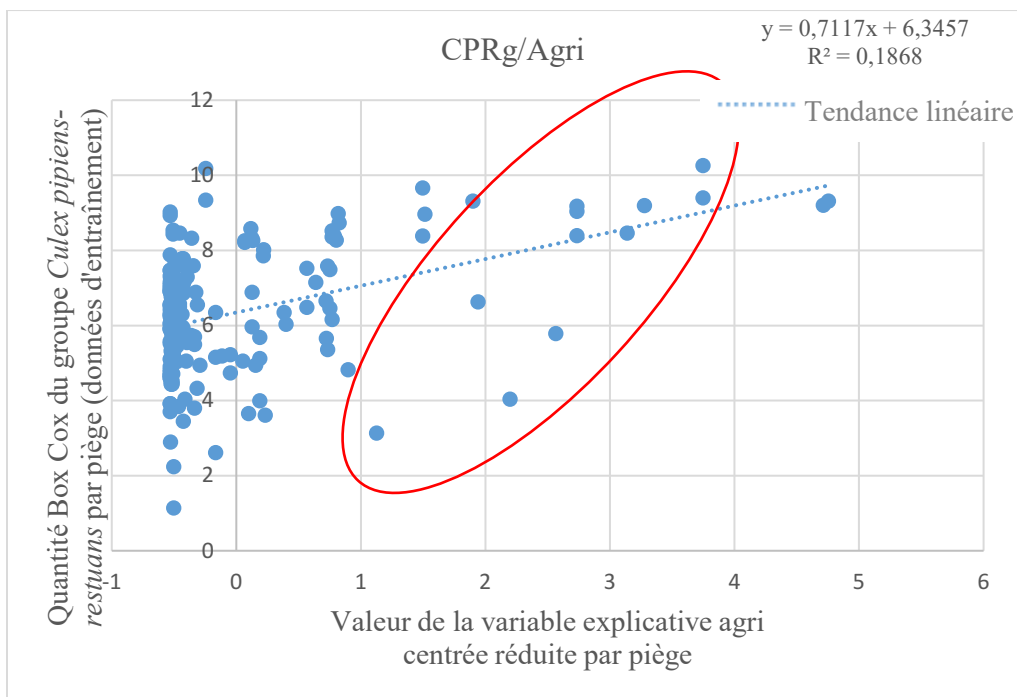


Figure 8 : Relation non linéaire entre la variable centrée réduite Agri et les quantités transformées de *Culex pipiens-restuans*.

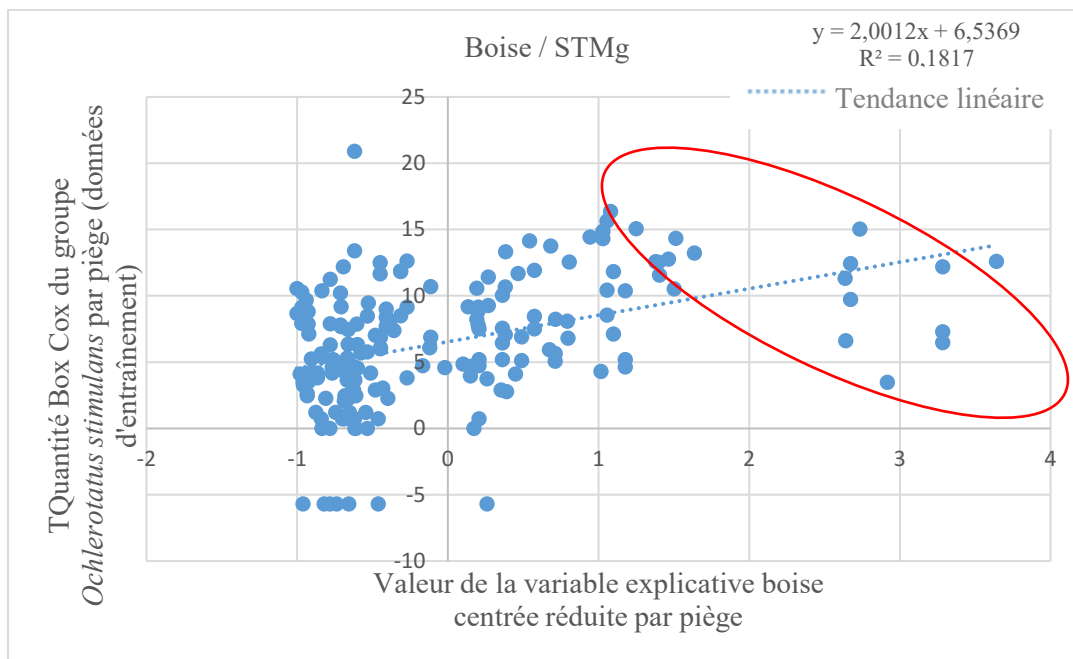


Figure 9 : Relation non linéaire entre la variable centrée réduite Boise et les quantités transformées de moustiques d'*Ochlerotatus stimulans*.

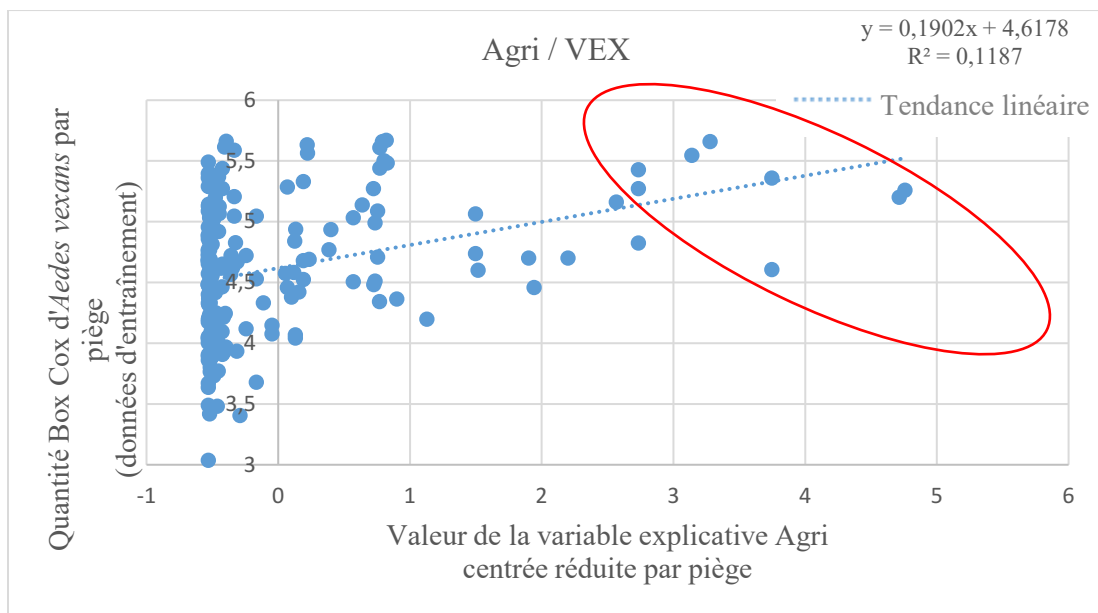


Figure 10 : Relation non linéaire entre la variable centrée réduite Agri et les quantités transformées d'Aedes vexans.

Après évaluation de la *p-value* des termes d'interaction et des termes quadratiques en régression univariée (tableau complet en annexe), il reste trois termes d'interaction significatifs pour CQP, quatre termes d'interaction et un terme quadratique significatifs pour VEX et deux termes d'interaction et un terme quadratique significatifs pour CPRg et STMg (tableau 24). Avec et sans ces nouvelles variables composées, nous obtenons six modèles par espèce (trois MLM1sans et trois MLM1avec), sauf quatre pour CPRg : un par seuil VIF, agrémenté ou non de termes d'interaction ou quadratique.

Tableau 24 : Variables simples et termes quadratiques et/ou d'interactions significatifs pour chaque espèce.

CPRg	CQP	STMg	VEX
Agri	Agri	Urbain	Eau
Prairie	Urbain	Prairie	Agri
Boise	Boise	Boise	Pfev
TNsep_1	Pjan	Pjan	Psep_1
Prairie x Boise	Psep	Pmar	Pjan
Agri x Agri	Pjul	Pjun	Pjul
Boise x TNsep_1	Paou	Paou	Psep
	Urbain x Boise	Psep	Agri x Agri
	Psep x Paou	Boise x Boise	Agri x Psep_1
	Pjul x Paou	Pjun x Paou	Psep_1 x Pjan
		Paou x Psep	Agri x Pjul
			Pjan x Psep

	VIF <=10
	VIF <=5
	VIF <=3

CPRg : groupe *Culex-piapiens restuans* – CQP : *Coquilleltidia perturbans*  
STMg : groupe *Ochlerotatus stimulans* – VEX : *Aedes vexans*.

Le résultat des MLM1avec3, 5 et 10 sont présentés dans les tableaux 25 à 29.

- Groupe *Culex pipiens-restuans*

Le MLM1avec3 donne le même résultat que le MLM1sans3, car aucune des variables composées n'est restée significative. Les paramètres du MLM1avec10 sont détaillés dans le tableau 25.

Tableau 25 : MLM1avec10 pour le groupe *Culex pipiens-restuans*.

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,369	0,117	57	54,415	< 0,0001	6,139	6,600
Agri	0,658	0,121	57	5,455	< 0,0001	0,420	0,896
Prairie	-0,422	0,121	57	-3,473	0,001	-0,662	-0,182
Boise	-0,256	0,128	57	-2,007	0,050	-0,507	-0,004
TNsep_1	-0,218	0,103	57	-2,103	0,040	-0,422	-0,013
Boise x TNsep_1	0,239	0,102	57	2,342	0,023	0,038	0,441

Il ne reste que le terme d'interaction entre Boise et TNsep\_1 qui soit significatif et la *p-value* de la variable de COS Boise, quoique toujours acceptable, est maintenant limite (0,05). Toutes les variables du modèle MLM1sans10 restent significatives.

- *Coquillettidia perturbans*

Aucun terme d'interaction ne reste significatif dans les MLM1. Comme il n'y avait pas de terme quadratique à tester, il n'y a donc pas de modèle additionnel pour CQP.

- Groupe *Ochlerotatus stimulans*

Tableau 26 : MLM1avec3 pour le groupe *Ochlerotatus stimulans*.

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,547	0,342	59	19,140	< 0,0001	5,872	7,222
Urbain	-1,527	0,482	59	-3,165	0,002	-2,479	-0,575
Boise	2,047	0,624	59	3,280	0,002	0,816	3,278
Boise x Boise	-1,285	0,522	59	-2,460	0,017	-2,316	-0,255

Tableau 27 : MLM1avec5 pour le groupe Ochlerotatus stimulans.

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,506	0,336	58	19,351	< 0,0001	5,843	7,170
Urbain	-1,532	0,474	58	-3,234	0,002	-2,467	-0,597
Boise	1,981	0,613	58	3,230	0,002	0,771	3,191
Pjan	-0,661	0,260	58	-2,540	0,014	-1,174	-0,147
Boise x Boise	-1,211	0,514	58	-2,358	0,022	-2,225	-0,198

Tableau 28 : MLM1avec10 pour le groupe Ochlerotatus stimulans.

Source	Valeur	Erreur standard	Degré de liberté	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	6,533	0,331	58	19,715	< 0,0001	5,879	7,187
Urbain	-1,448	0,467	58	-3,103	0,003	-2,369	-0,527
Boise	1,907	0,604	58	3,155	0,003	0,714	3,099
Pjun	-1,127	0,240	58	-4,693	< 0,0001	-1,601	-0,653
Boise x Boise	-1,146	0,507	58	-2,260	0,028	-2,147	-0,146

Dans les trois cas (tableaux 26 à 28), seul le terme quadratique Boise<sup>2</sup> reste significatif et toutes les variables en présence dans les MLM1 sont conservées dans les MLM1avec correspondants. L'ajout du terme quadratique Boisé<sup>2</sup> améliore la *p-value* et le coefficient de participation de la variable simple Boise dans chacun des trois modèles. Cependant, cela n'a qu'une influence modérée sur les autres variables simples.

- *Aedes vexans*

Il n'y a que le MLM1avec10 (tableau 29) qui conserve à la fois le terme quadratique Agri<sup>2</sup> et le terme d'interaction combinant les précipitations de janvier et celles de septembre. Les MLM1avec 3 et 5 rejettent toutes les variables composées pour revenir aux paramètres des MLM1sans correspondants.

Tableau 29 : MLM1avec10 pour Aedes vexans.

Source	Valeur	Erreur standard	DDL	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Constante	4,615	0,035	56	132,824	< 0,0001	4,546	4,683
Agri	0,256	0,071	56	3,584	0,001	0,115	0,396
Pfev	-0,126	0,037	56	-3,440	0,001	-0,198	-0,054
Pjan	0,227	0,036	56	6,344	< 0,0001	0,156	0,297
Psep	0,126	0,032	56	3,902	0,000	0,062	0,190
Agri x Agri	-0,158	0,069	56	-2,294	0,026	-0,294	-0,022
Pjan x Psep	0,159	0,039	56	4,091	0,000	0,082	0,236

## 5.2.5. Étape 5 : choix du meilleur modèle

- Groupe *Culex pipiens-restuans*

Tableau 30 : Évaluation des modèles pour le groupe *Culex pipiens-restuans*.

	Critère d'information Akaïke	Shapiro- Wilk	Anderson- Darling	Lilliefors	Jarque-Bera
MLM1sans3	651	0,091	0,090	0,261	0,154
MLM1sans10	652	0,167	0,120	0,147	0,187
MLM1avec10	651	0,202	0,193	<b>0,047</b>	0,260

Les scores AIC sont très semblables entre les trois modèles (tableau 30). Cependant, si l'on souhaite évaluer l'impact des changements climatiques à long terme, le modèle MLM1sans3 ne comporte pas de variables météorologiques, il ne serait pas utile en matière de prédiction future. Le résultat aux tests de normalité est en moyenne légèrement supérieur pour le MLM1avec10. C'est donc le MLM1avec10 (tableau 25) qui est sélectionné comme modèle final MLM1f. Sa formule se décline comme suit :

$$y = 6,369 + 0,658x \text{ Agri} - 0,422 x \text{ Prairie} - 0,256 x \text{ Boise} - 0,218 x \text{ TNsep\_1} + 0,239 x \text{ Boise} x \text{ TNsep\_1} \quad \text{Équation 1}$$

Les tests de normalité s'interprètent comme suit : le risque de rejeter l'hypothèse nulle  $H_0$  selon laquelle la variable dont provient l'échantillon suit une loi Normale alors qu'elle est vraie est de  $x$  %. Dans le tableau 30,  $x$  varie donc entre 4,7 % (test de Lilliefors du MLM1avec 10) et 26 % (test de Jarque-Bera du MLM1avec10). La valeur en gras indique que le test a échoué. C'est le cas lorsque la *p-value* calculée par XLSTAT associée au test est inférieure à 0,05 dans un intervalle de confiance de 95 % et qu'il faut alors rejeter  $H_0$ .

L'abondance de CPRg est donc prédite par un modèle linéaire combinant trois variables d'occupation du sol (l'agriculture, la prairie et le boisé sous formes simples), ainsi qu'une variable climatique (TNsept-1) et un terme d'interaction combinant la variable climatique avec le boisé. La classe agricole tout comme le terme d'interactions ont un effet positif sur la densité de moustiques : il serait tentant d'en conclure que, plus la couverture agricole est dense, plus il y aura de CPR. Cependant, toutes les variables explicatives composant les modèles sont centrées réduites. Ainsi, les valeurs brutes les plus faibles de chacune des variables sont transformées en une valeur centrée réduite négative, et inversement. Cela rend l'interprétation plus délicate

puisque les valeurs des données environnementales se répartissent de façon plus ou moins équivalente entre valeurs négatives et valeurs positives. Le coefficient de participation positif de la variable agricole devrait donc se traduire par le fait que les faibles densités de milieu agricole ont une influence négative sur les abondances de CPRg et qu'à partir d'un certain seuil de couverture agricole en augmentant, cette influence devient de plus en plus positive. Les autres variables ont toutes un impact négatif : plus il y a de prairies ou de boisés, moins il y aura de CPRg et plus les prairies ou boisés seront rares, plus cela aura un impact positif sur les CPRg. Pour les TNsep\_1, si ces températures centrées réduites sont négatives, elles auront un impact positif sur l'abondance des CPRg pour l'été de l'année suivante, et inversement. L'impact météorologique semble cependant moins élevé que celui des variables de COS, et en particulier que celui de l'agriculture.

- *Coquillettidia perturbans*

Tableau 31 : Évaluation des modèles pour *Coquillettidia perturbans*.

	Critère d'information Akaike	Shapiro- Wilk	Anderson- Darling	Lilliefors	Jarque-Bera
MLM1sans3	573	<b>0,048</b>	0,099	0,244	0,135
MLM1sans5	536	<b>0,010</b>	<b>0,003</b>	<b>0,004</b>	0,089
MLM1sans10	530	<b>0,021</b>	<b>0,014</b>	0,093	<b>0,037</b>

Pour CQP, c'est le MLM1sans10 (tableau 17) qui obtient le meilleur score AIC (tableau 31). En revanche, les résultats aux tests de normalité sont moins bons que ceux du MLM1sans3. Le MLM1sans10 échoue à trois tests (en gras dans le tableau) sur les quatre tests, alors que le MLM1sans3 n'échoue qu'à un seul test. Comme on cherche surtout à vérifier si le comportement des résidus est proche de la normale ou complètement erratique, l'évaluation des graphiques *Q-Q plot* et *P-P plot*, ainsi que la forme des histogrammes des résidus de Student validera ou non la pertinence de conserver le MLM1sans10 comme modèle final (MLM1f), que la nette différence de performance à l'AIC favorise. La formule qui résulte du MLM1f pour CQP est donc la suivante :

$$y = 3,7 - 0,535 \times \text{Urbain} + 0,244 \times \text{Pjan} + 0,265 \times \text{Paout} \quad \text{Équation 2}$$

La densité de CQP est donc prédite par une combinaison de deux variables climatiques, d'hiver et d'été (les précipitations moyennes des mois de janvier et août) et une variable d'occupation du sol (le milieu urbain). La variable de COS a un coefficient de participation négatif. Ainsi, plus la couverture urbaine est dense, moins il y aura de CQP, mais en deçà d'un certain pourcentage, correspondant à la moyenne de ces pourcentages avant qu'elles ne soient centrées réduites, la rareté du milieu urbain aura un impact positif sur les abondances de CQP. Les variables de précipitation ont au contraire un coefficient de participation positif : plus il y a de précipitations sous forme de pluie en août ou en janvier, sous forme de pluie ou de neige, plus il y aura de CQP au cours de l'été. À l'inverse, de faibles précipitations en janvier ou août auront une influence négative sur les abondances de CQP.

- Groupe *Ochlerotatus stimulans*

Tableau 32 : Évaluation des modèles pour le groupe *Ochlerotatus stimulans*.

	Critère d'information Akaïke	Shapiro- Wilk	Anderson- Darling	Lilliefors	Jarque-Bera
MLM1sans3	1035	<b>0,002</b>	<b>0,011</b>	0,157	<b>0,004</b>
MLM1avec3	1031	<b>0,007</b>	<b>0,028</b>	0,309	<b>0,003</b>
MLM1sans5	1031	<b>0,047</b>	0,106	0,401	0,052
MLM1avec5	1027	0,068	0,179	0,593	<b>0,033</b>
MLM1sans10	1017	0,567	0,494	0,518	0,273
MLM1avec10	1013	0,546	0,549	0,666	0,173

Pour STMg, on constate que c'est le MLM1avec10 qui obtient les meilleurs scores aux tests d'AIC et de normalité (tableau 32). C'est donc celui qui est sélectionné comme modèle final MLM1f. Ses paramètres sont présentés dans le tableau 28 et la formule associée est présentée avec l'équation 3. La qualité du test de normalité est ici délicate à différencier entre les modèles sans et avec termes d'interaction et/ou quadratique, mais la moyenne des résultats au test de normalité par modèle est supérieure pour celui avec termes d'interaction.

$$y = 6,533 - 1,448 \times \text{Urbain} + 1,907 \times \text{Boise} - 1,127 \times \text{Pjun} - 1,146 \times \text{Boise}^2 \quad \text{Équation 3}$$

Les variations d'abondance de STMg sont ainsi traduites par un modèle linéaire combinant deux variables d'occupation du sol (l'urbain, ainsi que le boisé sous forme simple et sous forme quadratique), et d'une seule variable climatique d'été (les précipitations moyennes du mois de

juin). La variable Boise a un effet positif sur la densité de moustiques alors que son terme quadratique a une influence négative : l'effet de la couverture boisée est donc délicat à interpréter en fonction des valeurs négatives ou positives de la variable centrée réduite. Les variables de COS Urbain et les précipitations totales de juin (Pjun) ont toutes deux un impact négatif : plus la densité urbaine est importante et plus il y a de précipitations à la fin du printemps, moins il y aura de STMg au cours de l'été. Et de faibles proportions de couverture urbaine ainsi que de faibles précipitations printanières auront un effet positif sur les quantités de STMg.

- *Aedes vexans*

Tableau 33 : Évaluation des modèles pour *Aedes vexans*.

	Critère d'information Akaïke	Shapiro- Wilk	Anderson- Darling	Lilliefors	Jarque-Bera
MLM1sans3	291	0,530	0,459	0,291	0,357
MLM1sans5	286	0,700	0,338	0,157	0,701
MLM1sans10	256	0,470	0,446	0,445	0,386
MLM1avec10	245	0,626	0,498	0,570	0,556

Pour l'espèce VEX, c'est le MLM1avec10 (tableau 29) qui performe le mieux à l'AIC (tableau 33). Les tests de normalité sont également meilleurs en moyenne pour ce modèle. Il est donc sélectionné comme étant le meilleur modèle final MLM1f. Ce modèle se traduit selon l'équation 4.

$$y = 4,615 + 0,256 \times \text{Agri} - 0,126 \times \text{Pfev} + 0,227 \times \text{Pjan} + 0,126 \times \text{Psep} - 0,158 \times \text{Agri}^2 + 0,159 \times \text{Pjan} \times \text{Psep} \quad \text{Équation 4}$$

La densité de VEX est ainsi prédite par un modèle linéaire combinant une variable d'occupation du sol (l'agriculture sous forme simple et sous forme quadratique), ainsi que trois variables climatiques d'hiver et d'automne (les précipitations moyennes des mois de janvier, février et septembre de l'année en cours) dont deux se combinent dans un terme d'interaction. La variable de COS a un effet positif sur la densité de moustiques : plus la couverture agricole est dense, plus il y aura de VEX tandis que les secteurs les moins agricoles auront un impact négatif sur les quantités de moustiques de cette espèce. Il en va de même avec les variables climatiques de précipitation en janvier et septembre. En revanche, la variable Pfev a un impact négatif : plus il y



a de précipitations en février, sous forme de pluie ou de neige, moins il y aura de VEX au cours de l'été suivant. Cependant, comme les précipitations de février et de septembre sont combinées dans un terme d'interaction doté d'un coefficient de participation positif, l'impact réel de ces précipitations sur l'abondance des VEX est plus difficile à interpréter.

Pour chaque espèce, ce sont finalement les modèles intégrant des variables météorologiques qui ont fourni les meilleurs résultats aux tests de normalité et d'AIC. Les formules mathématiques des quatre MLM1f (équations 1 à 4) serviront donc à générer la cartographie annuelle prédictive des densités de moustiques, à condition que ces modèles satisfassent aux critères de validation.

### 5.3. Phase de validation

#### 5.3.1. Étape 1 : validation interne

##### *Q-Q plots, P-P plots et homoscedasticité*

- Groupe *Culex pipiens-restuans*

Les *Q-Q plots* et *P-P plots* des résidus bruts du MLM1f (figure 11) confirment que leur distribution est proche de la normale. En effet, la répartition des résidus est très proche de la première bissectrice du graphique. La distribution des résidus de Student par rapport aux valeurs prédites (figure 12) est bien répartie entre la partie supérieure et inférieure et ne semble pas démontrer de structure particulière en dehors d'une tendance linéaire positive et de quelques points isolés. Cela laisse supposer que le modèle pourrait être amélioré, car toute la variabilité des abondances n'a pas su être traduite. Il est possible que le biais incontournable de la distribution non aléatoire des pièges et que les données météorologiques de faible résolution jouent un rôle dans ce résultat. La distribution normale des résidus bruts et l'homoscédasticité des résidus de Student pour CPRg sont visuellement jugées satisfaisantes dans le cadre de ce projet.

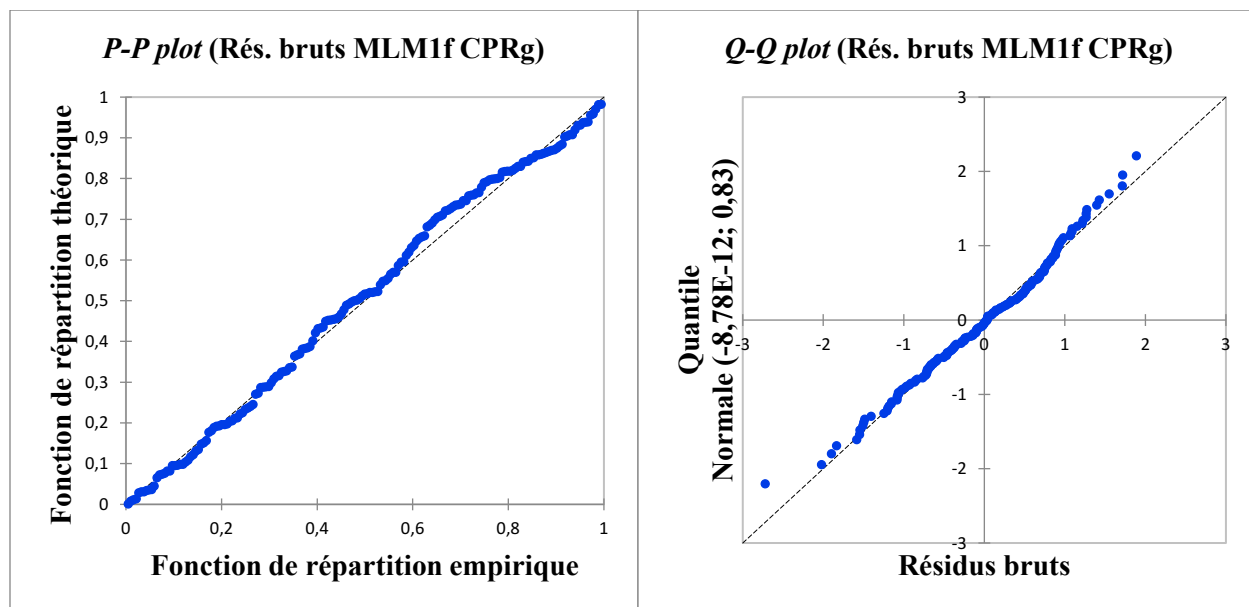


Figure 11 : P-P plot et Q-Q plot des résidus du meilleur modèle du groupe *Culex pipiens-restuans*.

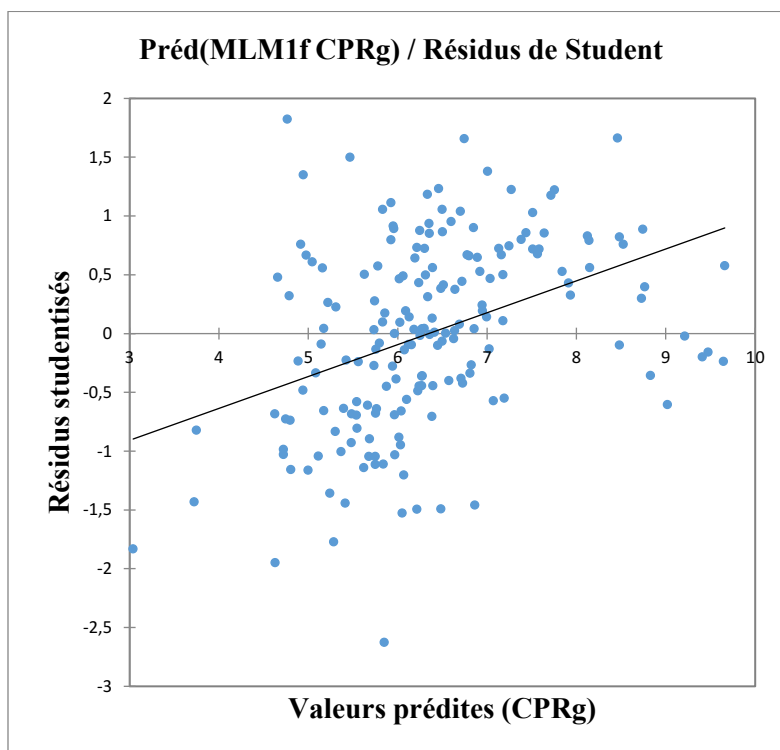


Figure 12 : Répartition des résidus de Student en fonction des prédictions du meilleur modèle du groupe *Culex pipiens-restuans*.

- *Coquillettidia perturbans*

Malgré le mauvais score aux tests de normalité, les *P-P plot* et *Q-Q plot* des résidus du MLM1f des CQP (figure 13) démontrent une distribution des résidus proche de la bissectrice et qui ressemble à celle de CPRg. Les défauts sont toutefois légèrement plus accentués pour CQP. L'homoscédasticité (figure 14) est illustrée par un nuage de points plutôt hétérogène, ce qui est recherché. Comme pour CPRg, on y retrouve toutefois une légère tendance linéaire positive.

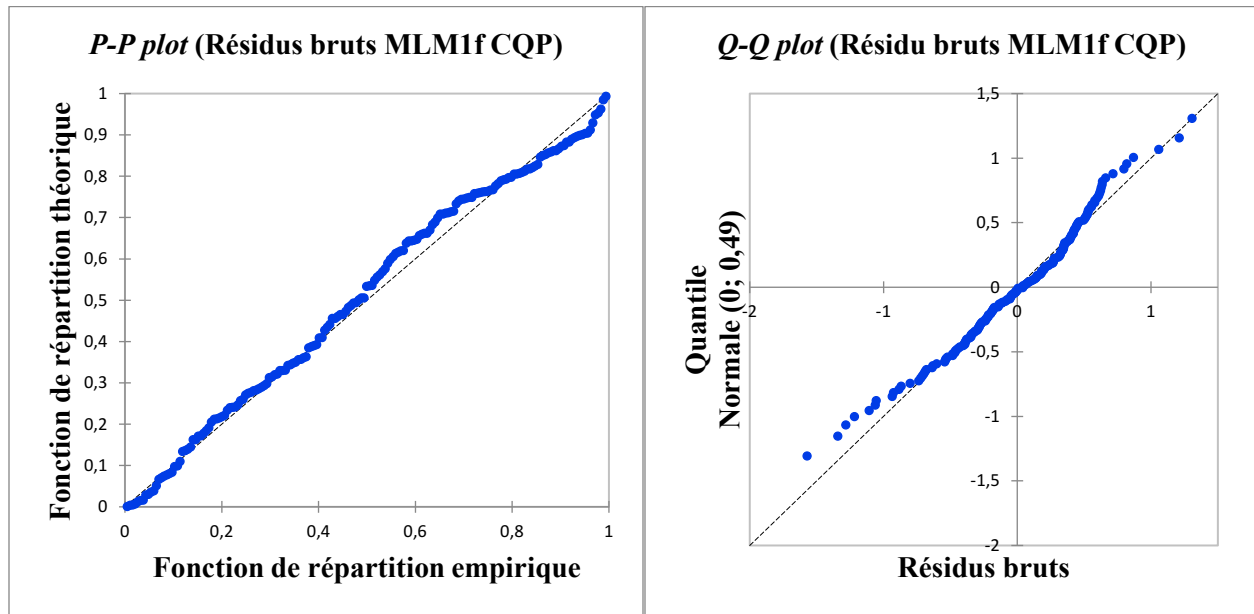


Figure 13 : P-P plot et Q-Q plot des résidus du meilleur modèle pour *Aedes vexans*.

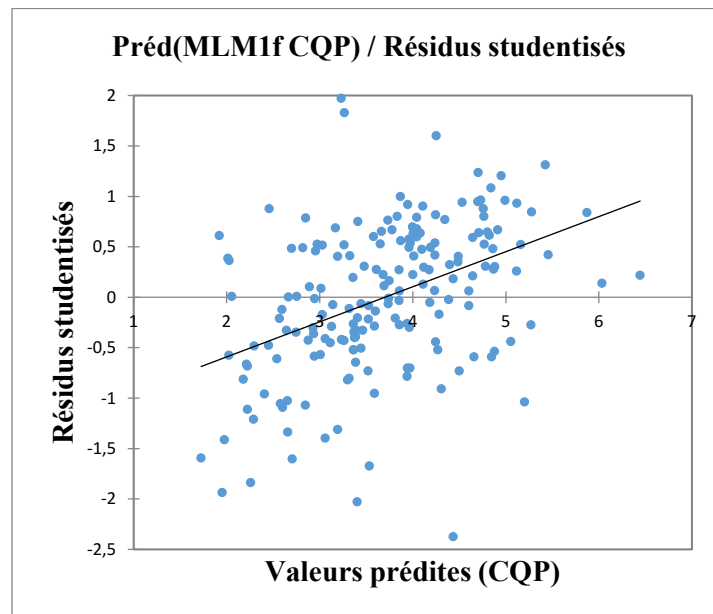


Figure 14 : Répartition des résidus de Student en fonction des prédictions du meilleur modèle pour *Coquillettidia perturbans*.

- Groupe *Ochlerotatus stimulans*

La distribution des résidus pour le *P-P plot* du MLM1f de STMg est ici presque collée à la bissectrice et le *Q-Q plot* ne montre que 3 valeurs extrêmes (figure 15), ce qui est excellent. Le nuage de points (figure 16) présente une forme homogène autour de l'axe d'origine, quelques points extrêmes et la même tendance linéaire positive pour CPRg et CQP.

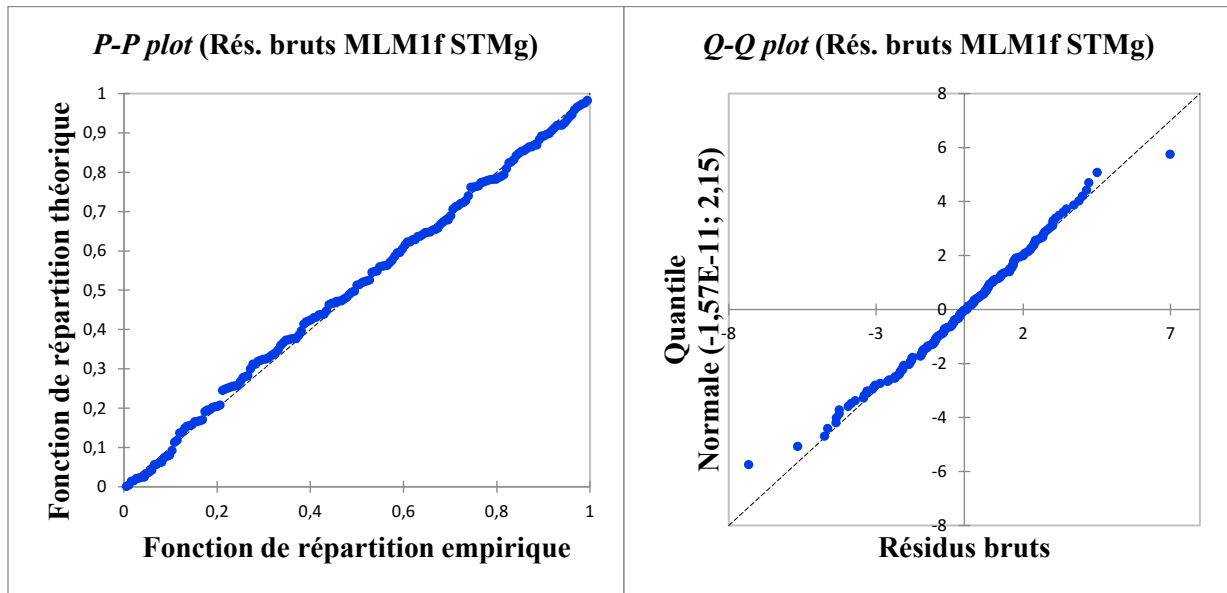


Figure 15 : P-P plot et Q-Q plot des résidus du meilleur modèle du groupe *Ochlerotatus stimulans*.

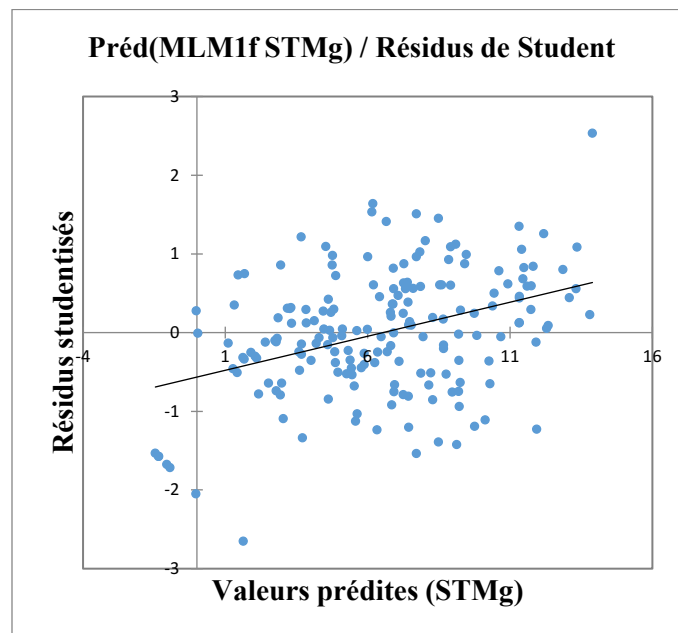


Figure 16 : Répartition des résidus de Student en fonction des prédictions du meilleur modèle du groupe *Ochlerotatus stimulans*.

- *Aedes vexans*

Les *P-P plot* et *Q-Q plot* des résidus bruts du MLM1f de VEX (figure 17) montrent une distribution proche de la normalité. La figure 18 montre un nuage de points sans structure particulière en dehors de la tendance linéaire positive semblable à celle des trois autres espèces, quoique moins prononcée. L'homoscédasticité des résidus de Student du modèle de VEX semble visuellement meilleure que pour CPRg, CQP ou STMg.

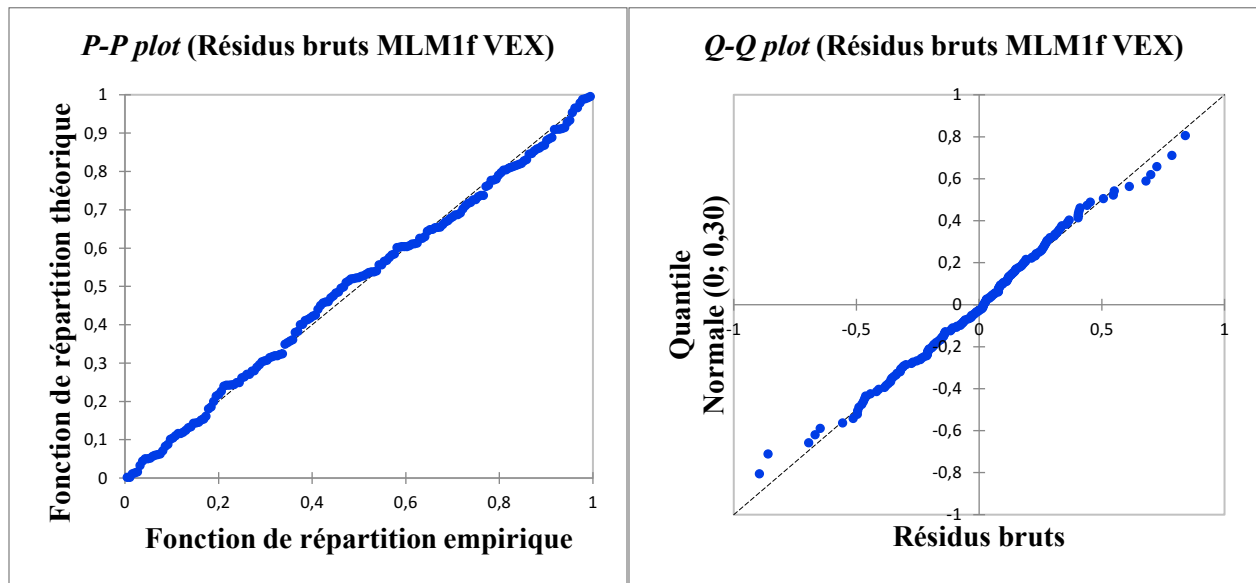


Figure 17 : P-P plot et Q-Q plot des résidus du meilleur modèle pour *Aedes vexans*.

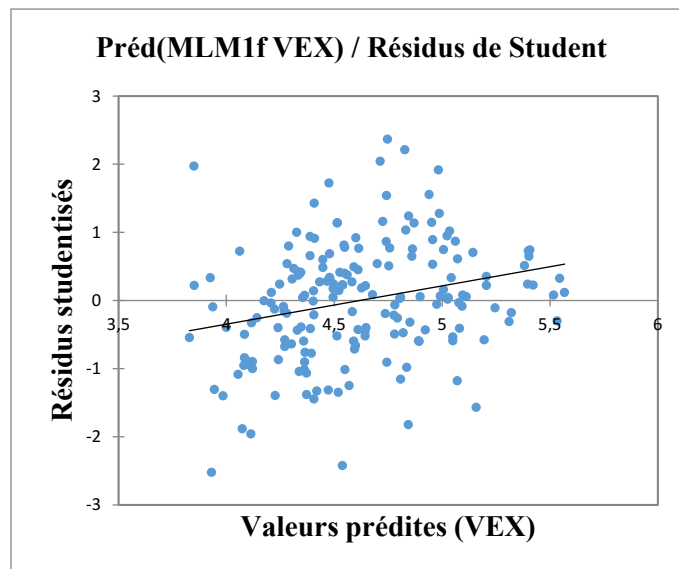


Figure 18 : Répartition des résidus de Student en fonction des prédictions du meilleur modèle pour *Aedes vexans*.

Les quatre espèces présentent donc une distribution de résidus bruts de leur MLM1f proche de la normale, ce qui est excellent. On observe toutefois, pour toutes les espèces, une tendance positive présente dans les nuages de points des résidus de Student en fonction des valeurs prédites. Cela renforce l'idée que la cause ne vient peut-être pas des modèles en eux-mêmes, mais plutôt de la qualité des données de base, puisque ce défaut est identique pour chacun des modèles.

### Histogrammes de distribution des résidus

- *Groupe Culex pipiens-restuans*

L'histogramme de distribution des résidus de Student du MLM1f de CPRg (figure 19) montre une distribution normale avec une moyenne quasi nulle ( $-3^{-12}$ ) et un écart-type est très proche de 1 (1,003). Cela reflète que les valeurs sont majoritairement centrées autour d'une moyenne nulle, ce qui est excellent puisque ces résidus ne sont pas centrés. XLSTAT n'arrondit pas les valeurs, ce qui donne une moyenne exacte de  $-3^{-12}$ , mais même avec quatre chiffres après la virgule, cela nous autoriserait à affirmer que la moyenne des résidus de Student est nulle. Leurs valeurs se répartissent dans un intervalle de  $[-3 ; +2,5]$ . Les statistiques descriptives associées montrent neuf valeurs hors intervalle, ce qui signifie que 95,1 % des valeurs sont situées entre -2 et +2, ce qui est recherché. Cela confirme la robustesse du MLM1f de CPRg.

- *Coquillettidia perturbans*

Les résidus de Student du MLM1f de CQP se répartissent normalement entre -3,5 et 3 avec une moyenne nulle et un écart-type de 1. Les valeurs à l'extérieur de l'intervalle cible comportent sept valeurs sous-estimées et 3 surestimées sur un total de 184 valeurs (figure 20). Avec 94,6 % de l'échantillonnage dans l'intervalle  $[-2 ; 2]$ , ce modèle est statistiquement robuste.

- *Groupe Ochlerotatus stimulans*

La moyenne des résidus de Student du MLM1f de STMg est égale à 0, son écart-type est égal à 1 et 95,7 % des valeurs se situent dans l'intervalle  $[-2 ; 2]$ . 8 valeurs sur 184 ont manqué leur cible, dont deux valeurs surestimées contre six sous-estimées (figure 21). La performance du modèle est donc excellente.

- *Aedes vexans*

L'histogramme des résidus de Student du MLM1f de VEX montre une distribution normale avec une moyenne nulle et un écart-type de 1. Les valeurs extrêmes sont réparties uniformément de part et d'autre de la courbe : cinq valeurs négatives contre six valeurs positives dans un intervalle absolu [2 ; 3] (figure 22). Il y a donc 94 % des valeurs dans la cible. Bien que légèrement inférieur aux 95 % recherchés, cela est suffisant pour poursuivre la phase cartographique avec ce modèle.

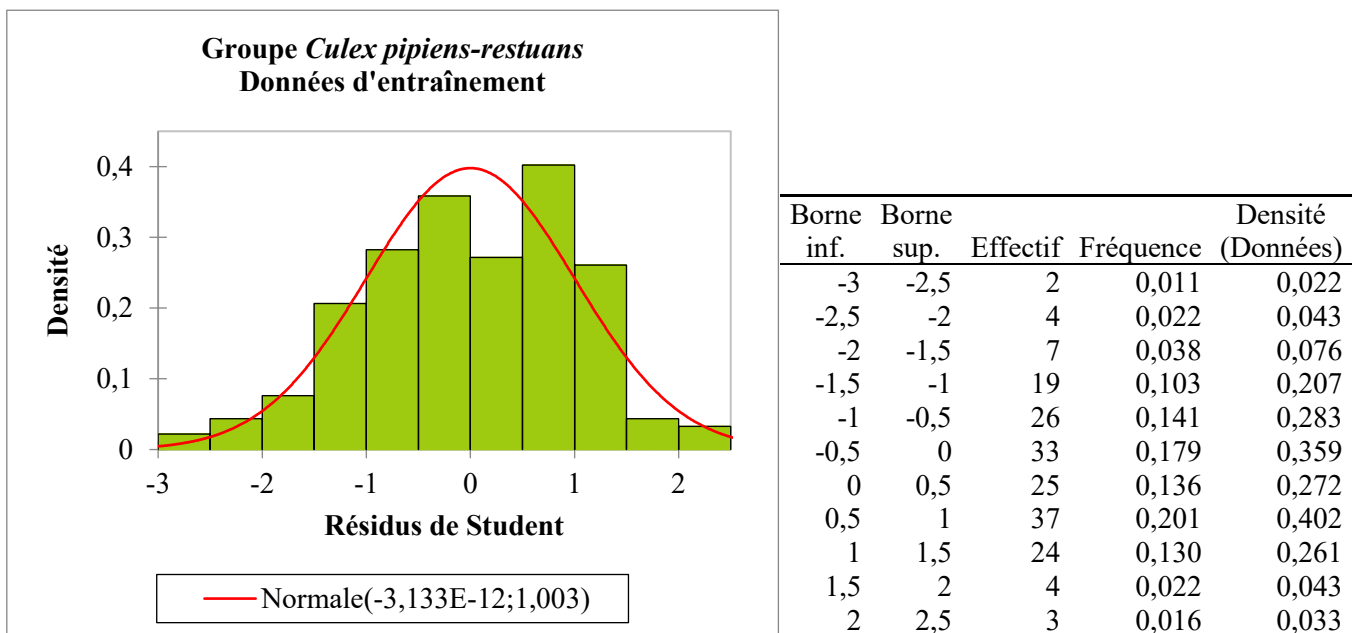


Figure 19 : Histogramme de distribution des résidus de Student du MLM1f pour le groupe *Culex pipiens-restuans* et statistiques descriptives associées (données d'entraînement).

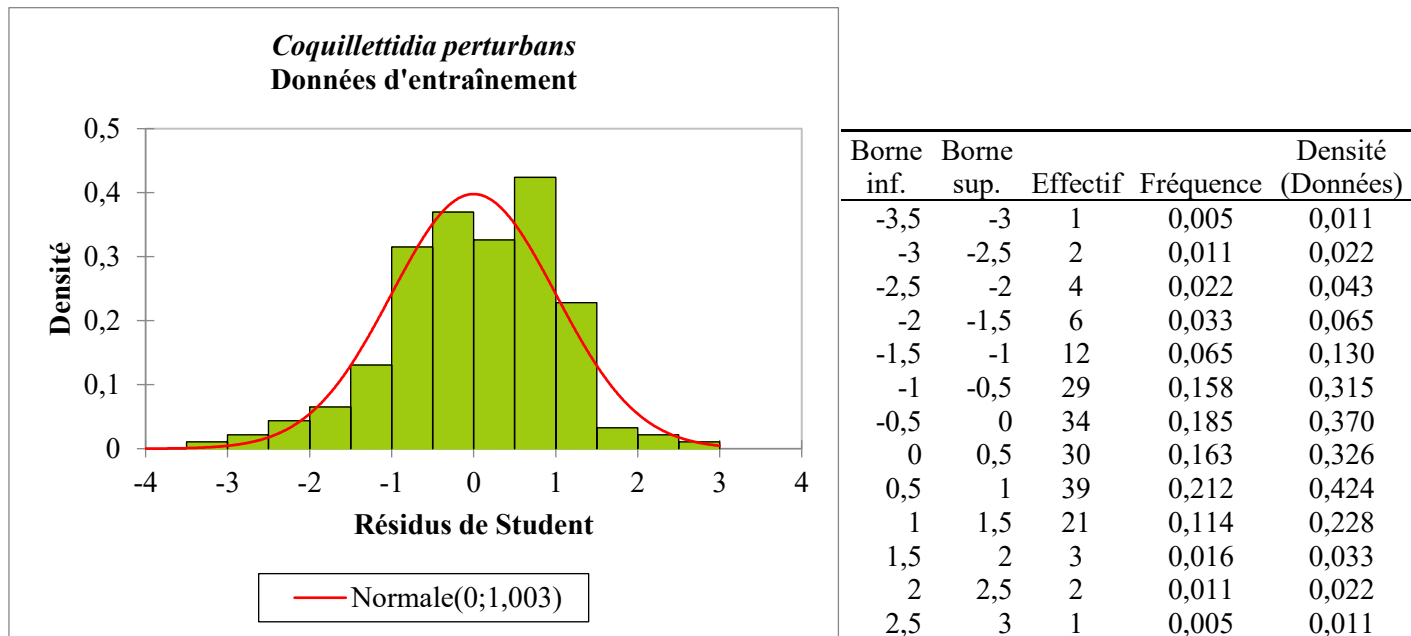


Figure 20 : Histogramme de distribution des résidus de Student du MLMlf pour *Coquillettidia perturbans* et statistiques descriptives associées (données d'entraînement).

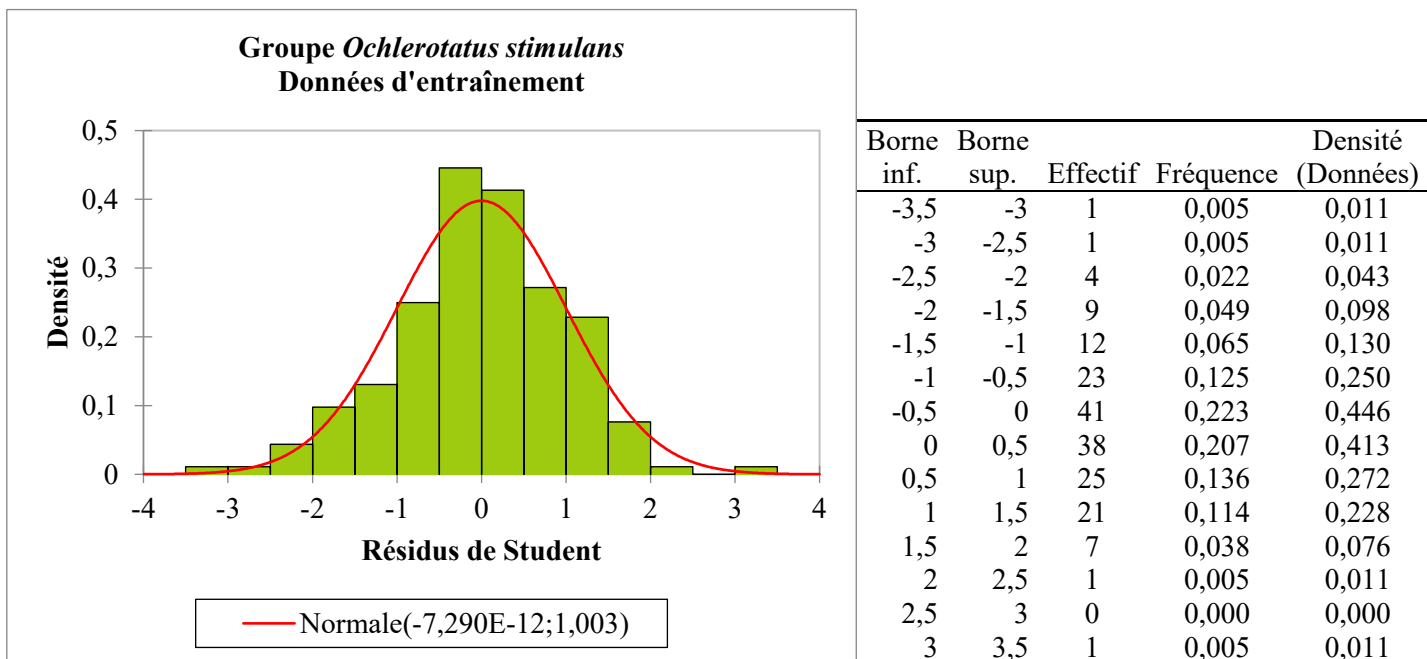


Figure 21 : Histogrammes de distribution des résidus de Student du MLMlf pour le groupe *Ochlerotatus stimulans* et statistiques descriptives associées (données d'entraînement).



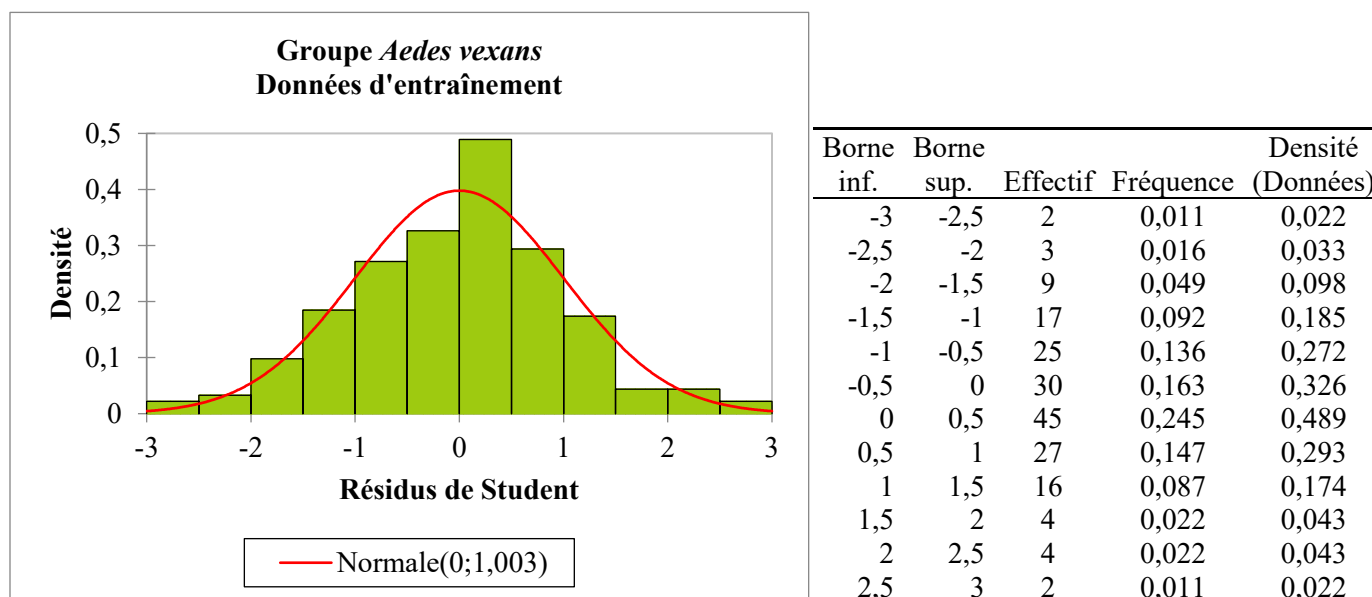


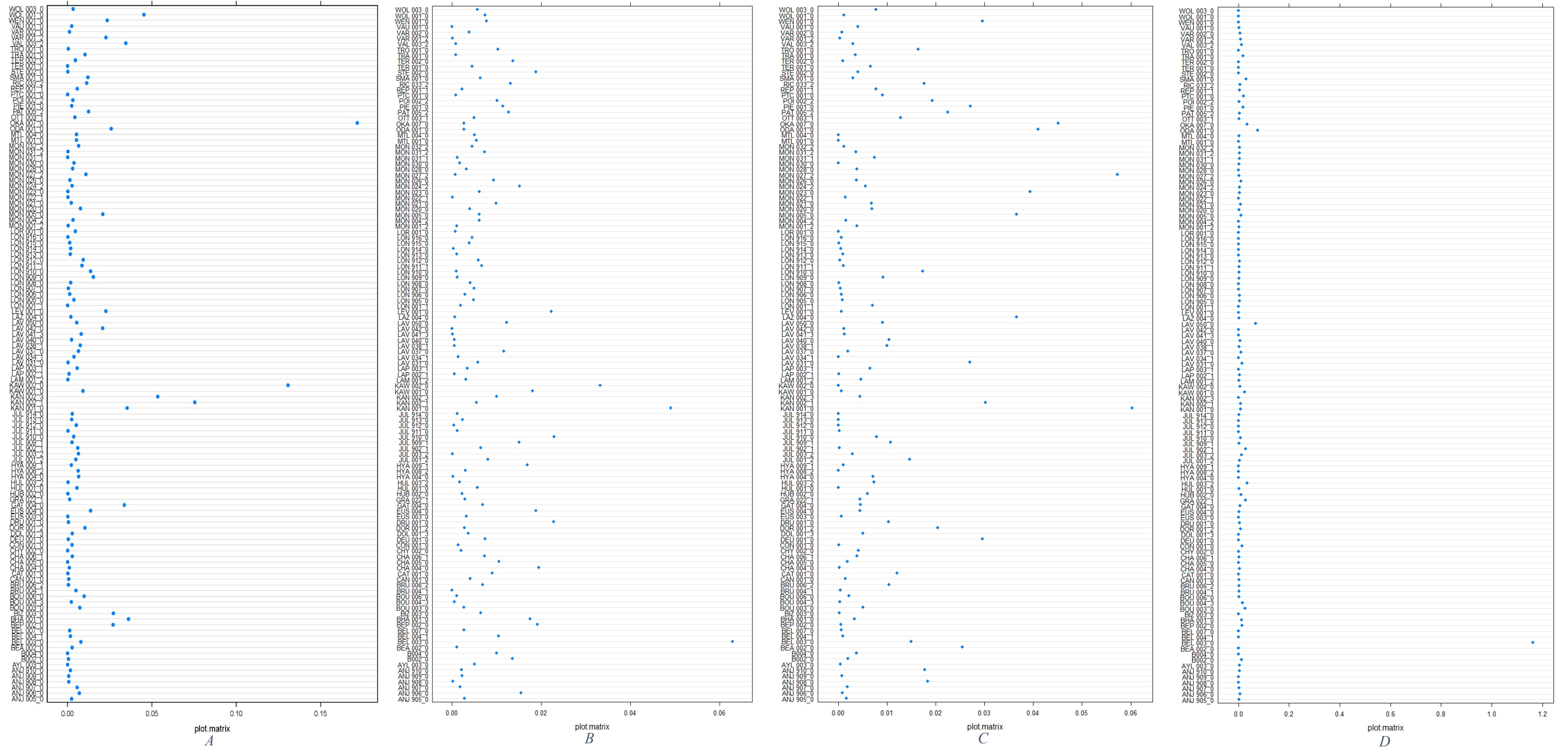
Figure 22 : Histogramme de distribution des résidus de Student du MLM1f pour *Aedes vexans* et statistiques descriptives associées (données d'entraînement).

Pour toutes les espèces, les moyennes sont nulles, environ 95 % des valeurs se situent dans l'intervalle  $[-2 ; 2]$  et les quelques valeurs hors intervalle se répartissent en deçà et au-delà des valeurs seuil selon une courbe normale. Cela indique que le MLM1f de chaque espèce est statistiquement robuste et qu'aucun des modèles ne surestime ni ne sous-estime globalement les abondances de moustiques comparativement aux données de capture. Cette qualité est donc un premier gage de confiance pour la cartographie à venir.

### Distance de Cook

La distance de Cook pour chacune des espèces confirme les bons résultats obtenus avec les histogrammes de résidus de Student (figures 23A à D). En effet, toutes les valeurs de toutes les distances de Cook sont inférieures à 1, en dehors d'un unique extrême légèrement supérieur à 1 pour VEX (pour le piège BEL 003\_0). Les distances sont toutes inférieures à 0,1 pour les trois autres espèces, sauf pour CPRg qui présente une seule valeur légèrement supérieure à 0,15 (pour le piège OKA 007\_0). La différence de graduation des abscisses donne ainsi au graphique de VEX une fausse impression de meilleur résultat que pour les autres espèces.

Compte tenu du faible score global à la distance de Cook pour toutes les espèces, nous considérons qu'il n'y a pas de valeurs aberrantes et qu'il n'est donc pas nécessaire de relancer la modélisation sans valeurs extrêmes.



Figures 23A à D : Distances de Cook des résidus de Student pour le meilleur modèle pour le groupe *Culex pipiens-restuans* (A), *Coquillettidia perturbans* (B), le groupe *Ochlerotatus stimulans* (C) et *Aedes vexans* (D).

### 5.3.2. Étape 2 : validation externe

- *Groupe Culex pipiens-restuans*

Le MLM1f de CPRg présente une moyenne de +0,602 dans un intervalle de [-2 ; +4,5] et dix valeurs extrêmes, toutes situées entre +2 et +4,5 (figure 24). 95,1 % des valeurs sont situées dans l'intervalle cible [-2 ; +2]. La courbe normale n'est donc pas centrée sur 0 et présente un décalage positif sur le graphique. Cela signifie que le modèle prédit globalement des valeurs de densité plus élevées que les valeurs observées, mais sa précision, validée par un fort taux de valeurs à l'intérieur de l'intervalle cible, ainsi que sa constance dans la surestimation validée par la normalité de la courbe, en font un excellent modèle prédictif.

- *Coquillettidia perturbans*

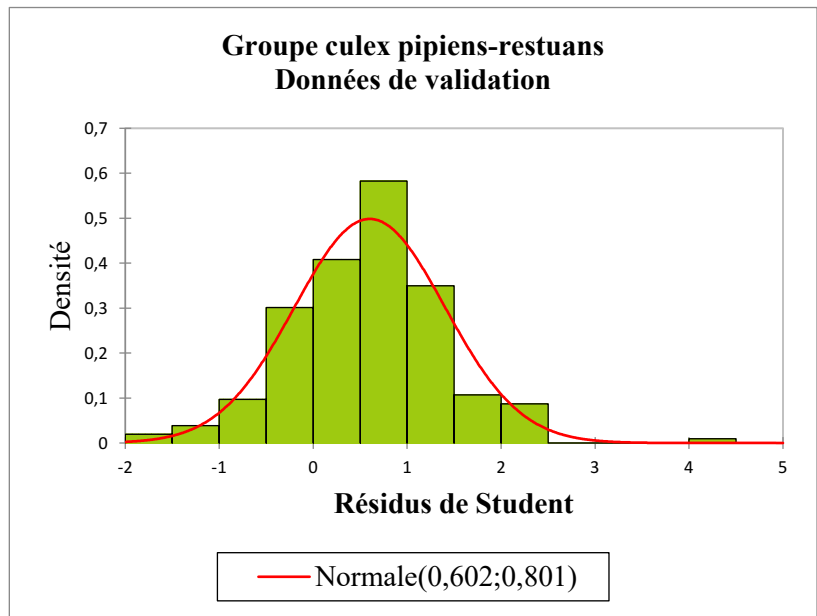
La moyenne du MLM1f de CQP se situe à -0,67 et les résidus se répartissent entre -4,5 et 1. Les valeurs extrêmes représentent 3,4 % de l'échantillonnage, soit sept valeurs hors intervalle (figure 25). Le modèle sous-estime donc légèrement les valeurs observées, mais de façon globalement constante. Il représente donc également un excellent modèle prédictif.

- *Groupe Ochlerotatus stimulans*

La moyenne du MLM1f de STMg est égale à -0,778 et 100 % des valeurs se situent dans l'intervalle [-2 ; 1,5] (figure 26). Bien que cela démontre l'excellente qualité du modèle et son potentiel prédictif performant, la moyenne négative indique que les valeurs calculées sont généralement sous-estimées.

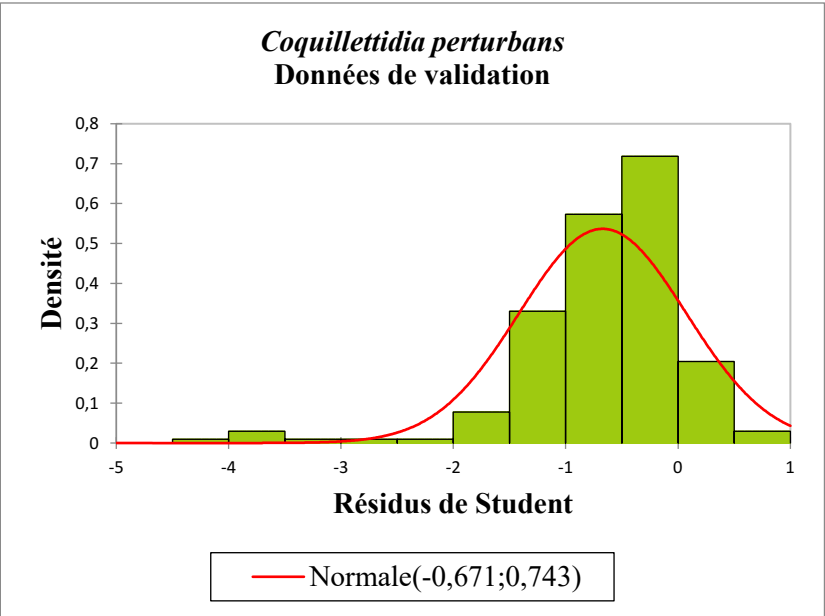
- *Aedes vexans*

L'histogramme du MLM1f de VEX présente une moyenne à -0,57. Les valeurs qui sortent de l'intervalle [-2 ; 2] sont toutes négatives et situées entre -3 et -2. Il y a trois valeurs sur 206 hors intervalle, soit 3,9 %, ce qui est inférieur à la limite de 5 % recherchée (figure 27). Ces résultats démontrent que le modèle est un excellent modèle prédictif, mais qu'il sous-estime les valeurs observées.



Borne inf.	Borne sup.	Effectif	Fréquence	Densité (Données)
-2	-1,5	2	0,010	0,019
-1,5	-1	4	0,019	0,039
-1	-0,5	10	0,049	0,097
-0,5	0	31	0,150	0,301
0	0,5	42	0,204	0,408
0,5	1	60	0,291	0,583
1	1,5	36	0,175	0,350
1,5	2	11	0,053	0,107
2	2,5	9	0,044	0,087
2,5	3	0	0,000	0,000
3	3,5	0	0,000	0,000
3,5	4	0	0,000	0,000
4	4,5	1	0,005	0,010

Figure 24 : Histogramme de distribution des résidus de Student du MLMIf pour le groupe *Culex pipiens-restuans* appliqué aux données de validation et statistiques descriptives associées.



Borne inf.	Borne sup.	Effectif	Fréquence	Densité (Données)
-4,5	-4	1	0,005	0,010
-4	-3,5	3	0,015	0,029
-3,5	-3	1	0,005	0,010
-3	-2,5	1	0,005	0,010
-2,5	-2	1	0,005	0,010
-2	-1,5	8	0,039	0,078
-1,5	-1	34	0,165	0,330
-1	-0,5	59	0,286	0,573
-0,5	0	74	0,359	0,718
0	0,5	21	0,102	0,204
0,5	1	3	0,015	0,029

Figure 25 : Histogramme de distribution des résidus de Student du MLMIf pour *Coquillettidia perturbans* appliqué aux données de validation et statistiques descriptives associées.

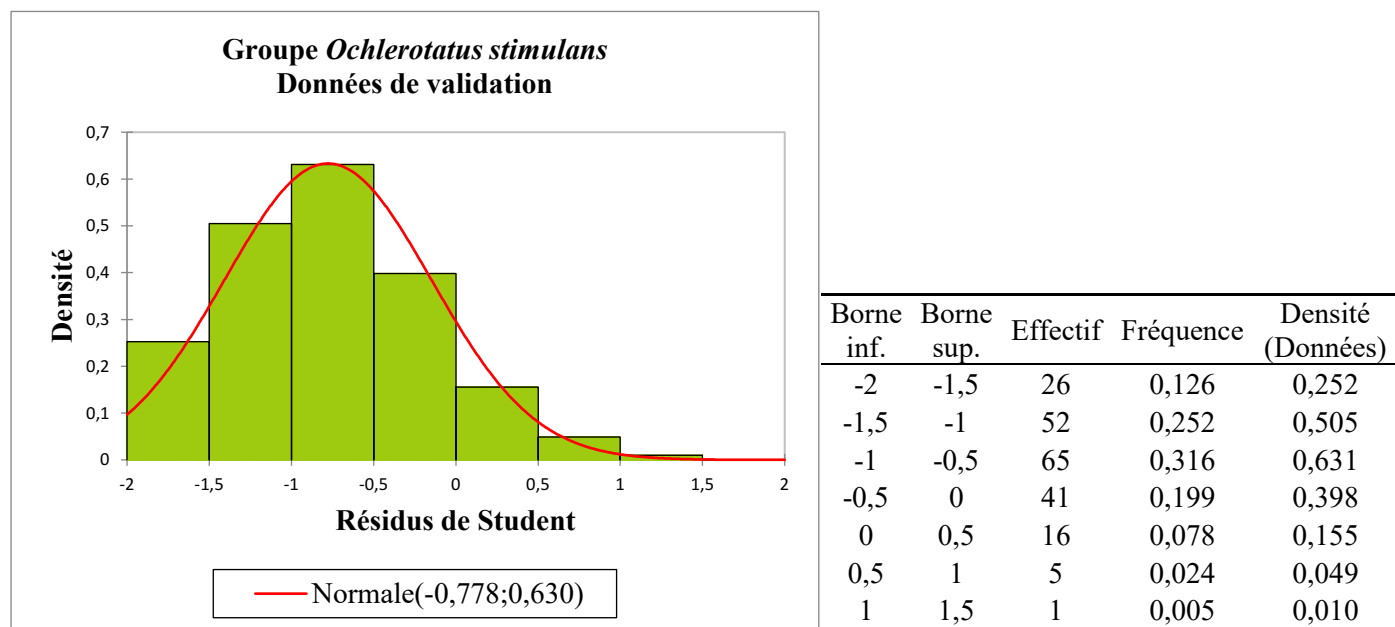


Figure 26 : Histogramme de distribution des résidus de Student du MLMIf pour le groupe *Ochlerotatus stimulans* appliqué aux données de validation et statistiques descriptives associées.

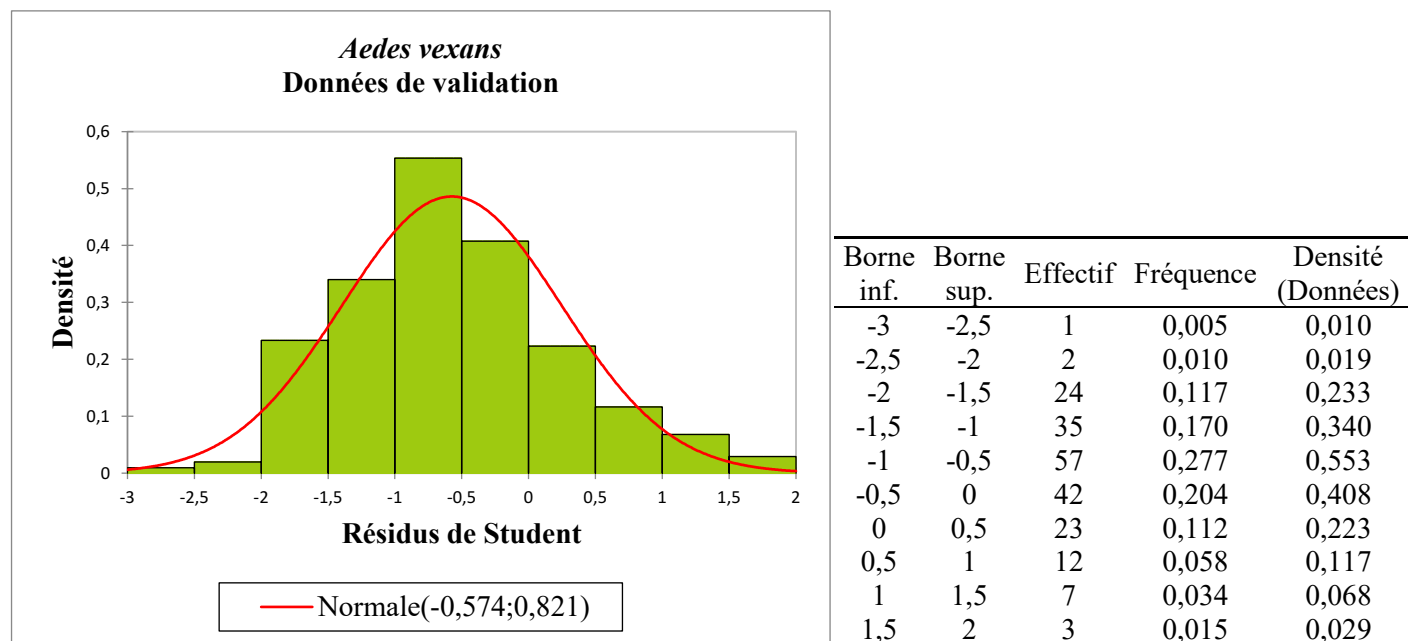


Figure 27 : Histogramme de distribution des résidus de Student du MLMIf pour *Aedes vexans* appliqué aux données de validation et statistiques descriptives associées.

Les histogrammes résultant de la distribution des résidus de Student des MLM1f appliqués aux données environnementales de validation nous informent donc que tous les modèles suivent une distribution normale et qu'ils ont une précision globale de prédiction supérieure à 95 % pour toutes les espèces, avec une tendance à la surestimation pour CPRg et à la sous-estimation pour CQP, STMg et VEX. La surestimation et la sous-estimation ne sont pas inquiétantes puisqu'on ne cherche pas à quantifier l'abondance de moustiques de façon exacte. En effet, toutes les moyennes sont décentrées d'une valeur inférieure à 1 et la distribution est régulière, sans pic anormal, en particulier dans les extrémités de la courbe. Ces modèles nous permettent donc de prédire avec confiance et précision les variations d'abondance capturées par les pièges en 2014. Les modèles entraînés avec les années 2003, 2004, 2005 et 2013 remplissent donc correctement leur rôle prédictif pour une année plus récente.

## 5.5. Phase cartographique

### 5.5.1. Étape 1 : carte des résidus des modèles

Relativement aux valeurs des résidus de Student des différents MLM1f de chaque espèce, la symbologie des cartes (figures 28 à 31) correspond à excellent pour l'intervalle  $[0 ; 1]$ , très bon pour  $[1 ; 2]$  et mauvais pour les valeurs de l'intervalle  $[2 ; 3]$  et très mauvais pour  $[3 ; 4,5]$ . Les intervalles négatifs sont classés de la même manière, selon les mêmes seuils.

Un regroupement ou une tendance particulière de la répartition des valeurs extrêmes existantes indiquerait que le modèle n'a pas capturé une ou des variables pouvant influencer les variations annuelles de densité de moustiques.

Les figures A illustrent la répartition des résidus de Student des données d'entraînement des MLM1f de chaque espèce. Cette répartition s'étend de Gatineau à Québec. Quant aux figures B, elles superposent les résidus de Student des données d'entraînement avec ceux des données de validation qui se concentrent sur la grande région de Montréal. Nous constatons, pour les quatre espèces, que le modèle appliqué aux données de validation démontre un ajustement spatial, c'est-à-dire dans les zones où il n'existait que peu de données d'entraînement, comme aux alentours de Saint-Jean-sur-Richelieu par exemple.

- *Groupe Culex pipiens-restuans*

Les valeurs extrêmes (en orange) du MLM1f généré avec les données d'entraînement sont réparties sans groupement sur le territoire (figure 28A). Elles se situent aux abords du fleuve St-Laurent et de la rivière Richelieu.

Les résidus issus de la validation externe superposés à ceux des données d'entraînement ne semblent pas confirmer ce schéma puisqu'on retrouve un « plus rouge » et un « plus orange » en plein cœur de l'île de Montréal. La superposition des deux jeux de résidus montre un seul groupement de valeurs négatives autour de la réserve faunique de Marguerite d'Youville (figure 28B). En effet, les pièges CHA006\_1, KAW001\_0, KAW 002\_0 des données d'entraînement et le piège CHA003\_0 de 2014 donnent des valeurs de captures supérieures à celles prédites par le MLM1f, alors que ce modèle a tendance à surestimer toutes les autres valeurs. Même si les valeurs des résidus de Student associées s'inscrivent dans l'intervalle  $[-2 ; 2]$ , ce groupement négatif en périphérie du refuge est remarquable.

- *Coquillettidia perturbans*

Les valeurs positives et négatives des résidus du MLM1f de CQP utilisant les données d'entraînement se répartissent uniformément sur le territoire, sauf autour du refuge faunique de Marguerite d'Youville où les valeurs sont sous-estimées (figure 29A). Les données de validation soulignent fortement cette disparité dans la distribution (figure 29B). En effet, tous les « moins rouges » ainsi qu'un « moins orange » sont concentrés sur cette zone. Les pièges correspondants sont CHA001\_0, CHA002\_0, CHA003\_0, CHA006\_0, CHA007\_0 et CHA008\_0. Les valeurs observées transformées Box Cox y sont très nettement supérieures à la moyenne (tableau 35).

- *Groupe Ochlerotatus stimulans*

La répartition des résidus de Student du MLM1f de STMG ne semble présenter ni tendance ni regroupement particulier et c'est exactement ce que l'on recherche (figure 30A). Quant aux résidus issus des données de validation, ils sont tous à l'intérieur de l'intervalle cible avec une tendance à la sous-estimation. La cartographie illustre ce résultat (figure 30B). On retrouve le groupement négatif autour du refuge faunique également, mais il n'est pas significatif comparé au reste du territoire pour STMg. Pourtant, l'observation des valeurs observées dans les six pièges CHA montre que les captures étaient en moyenne plus nombreuses qu'ailleurs (tableau 36). Le MLM1f de STMg semble donc être suffisamment performant pour prédire les pics d'abondance, même isolés.

- *Aedes vexans*

La figure 31A correspondant au MLM1f de VEX pour les données d'entraînement ne présente ni groupement ni tendance spatiale, ce qui est très bon. Concernant les données de validation, les « plus » et les « moins » verts se mêlent aux valeurs de l'intervalle  $[-1 ; 1]$  (figure 31B) ce qui confirme l'absence de groupement ou de tendance déjà observée. On retrouve toutefois la même sous-estimation dans les pièges CHA que pour les autres espèces. Les deux valeurs hors intervalle cible proviennent des pièges CHA001\_0 et CHA003\_0 (tableau 37).

Pour toutes les espèces, autant pour les données d'entraînement seules (figures 28A à 31A) que pour les données de validation superposées aux données d'entraînement (figures 28B à 31B), la distribution des valeurs classées « excellent » ainsi que celle des plus et des moins dans l'intervalle « très bon » sur le territoire est uniforme. Le groupement des valeurs sous-estimées,



situé autour de la réserve faunique de Marguerite d'Youville, est présent dans la cartographie de 2014, et ce, à différents degrés pour les quatre espèces. Cela veut dire que, systématiquement, le nombre de moustiques capturés par ces pièges est supérieur au nombre de moustiques estimés par les MLM1f. En effet, les observations enregistrées à cet endroit (tableaux 34 à 37) indiquent des moyennes de capture plus élevées que pour l'ensemble des autres valeurs de capture, et de façon beaucoup plus prononcée pour CQP (16,65 observations vs 5,63 prédites) et STMg (12,61 observations vs 5,97 prédites). Pour CPRg, si on regarde le seul piège qui présente un résidu inférieur à -1 (CHA 003\_0), alors on constate que le modèle a répondu avec une valeur prédite de 6,98 contre 6,58 en moyenne, soit une hausse de 5,7 %, mais insuffisante comparée au pic de 34,5 % des valeurs observées dans ce piège. Il s'agit donc vraisemblablement de la conséquence d'un phénomène local lié à l'environnement immédiat des pièges concernés et dont la teneur n'a pas été entièrement prise en compte dans les données environnementales que nous avons sélectionnées au départ. Cependant, si l'on observe maintenant les moyennes des valeurs prédites pour les six pièges CHA comparativement à la moyenne de l'ensemble des pièges, en dehors du MLM1f de VEX qui ne fait aucune différence, les trois autres MLM1f ont malgré tout correctement prédit un pic d'abondance pour ces pièges. En effet, CPRg enregistre une hausse de 1,2 % dans la moyenne de ces six pièges comparativement à la moyenne de tous les pièges, CQP une hausse de 17,3 % et STMg de 65,6 %. À ce titre, le MLM1f de VEX est le seul qui passe à côté du pic particulier. Comme les autres espèces, le modèle de VEX prédit toutefois adéquatement les variations d'abondances sur le territoire. L'ensemble des résultats de validation interne et externe valident donc la fiabilité spatiale de la capacité prédictive des MLM1f, et ce, pour toutes les espèces.

*Tableau 34 : Comparaison des valeurs du MLM1f du groupe Culex pipiens-restuans pour les pièges du secteur de la réserve faunique de Marguerite d'Youville par rapport à la moyenne de tous les pièges.*

Pièges	Valeurs observées Box Cox	Valeurs prédites Box Cox	Résidus bruts	Résidus de Student
CHA001_0	7,61	6,16	-1,45	-0,92
CHA002_0	6,31	6,47	0,16	0,10
<b>CHA003_0</b>	<b>8,60</b>	<b>6,98</b>	<b>-1,62</b>	<b>-1,03</b>
CHA006_0	6,32	7,01	0,69	0,44
CHA007_0	7,46	6,64	-0,83	-0,53
CHA008_0	5,22	6,69	1,47	0,93
<b>Moyennes des 6 pièges CHA</b>	6,92	6,66	-0,26	-0,17
<b>Moyennes des 206 pièges</b>	<b>5,63</b>	<b>6,58</b>	<b>0,95</b>	<b>0,60</b>

Tableau 35 : Comparaison des valeurs du MLMIf de *Coquillettidia perturbans* pour les pièges du secteur de la réserve faunique de Marguerite d'Youville par rapport à la moyenne de tous les pièges.

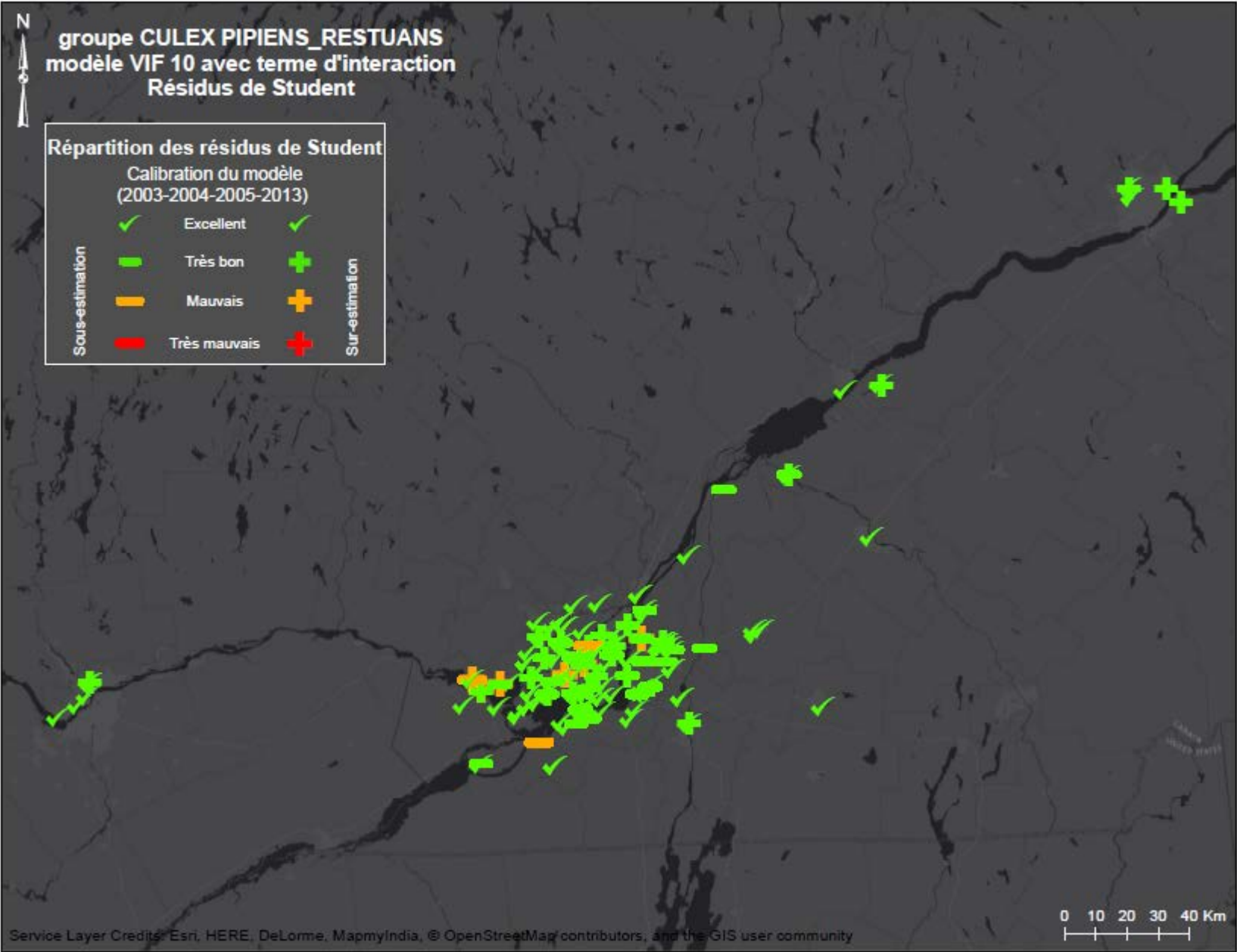
Pièges	Valeurs observées Box Cox	Valeurs prédites Box Cox	Résidus bruts	Résidus de Student
CHA001_0	19,00	3,87	-15,12	-4,36
CHA002_0	13,75	3,35	-10,39	-3,00
CHA003_0	17,72	3,98	-13,74	-3,96
CHA006_0	17,86	4,44	-13,43	-3,87
CHA007_0	16,86	3,70	-13,16	-3,80
CHA008_0	14,71	3,88	-10,83	-3,13
<b>Moyennes des 6 pièges CHA</b>	16,65	3,87	-12,78	-3,69
<b>Moyennes des 206 pièges</b>	<b>5,52</b>	<b>3,20</b>	<b>-2,33</b>	<b>-0,67</b>

Tableau 36 : Comparaison des valeurs du MLMIf du groupe *Ochlerotatus stimulans* pour les pièges du secteur de la réserve faunique de Marguerite d'Youville par rapport à la moyenne de tous les pièges.

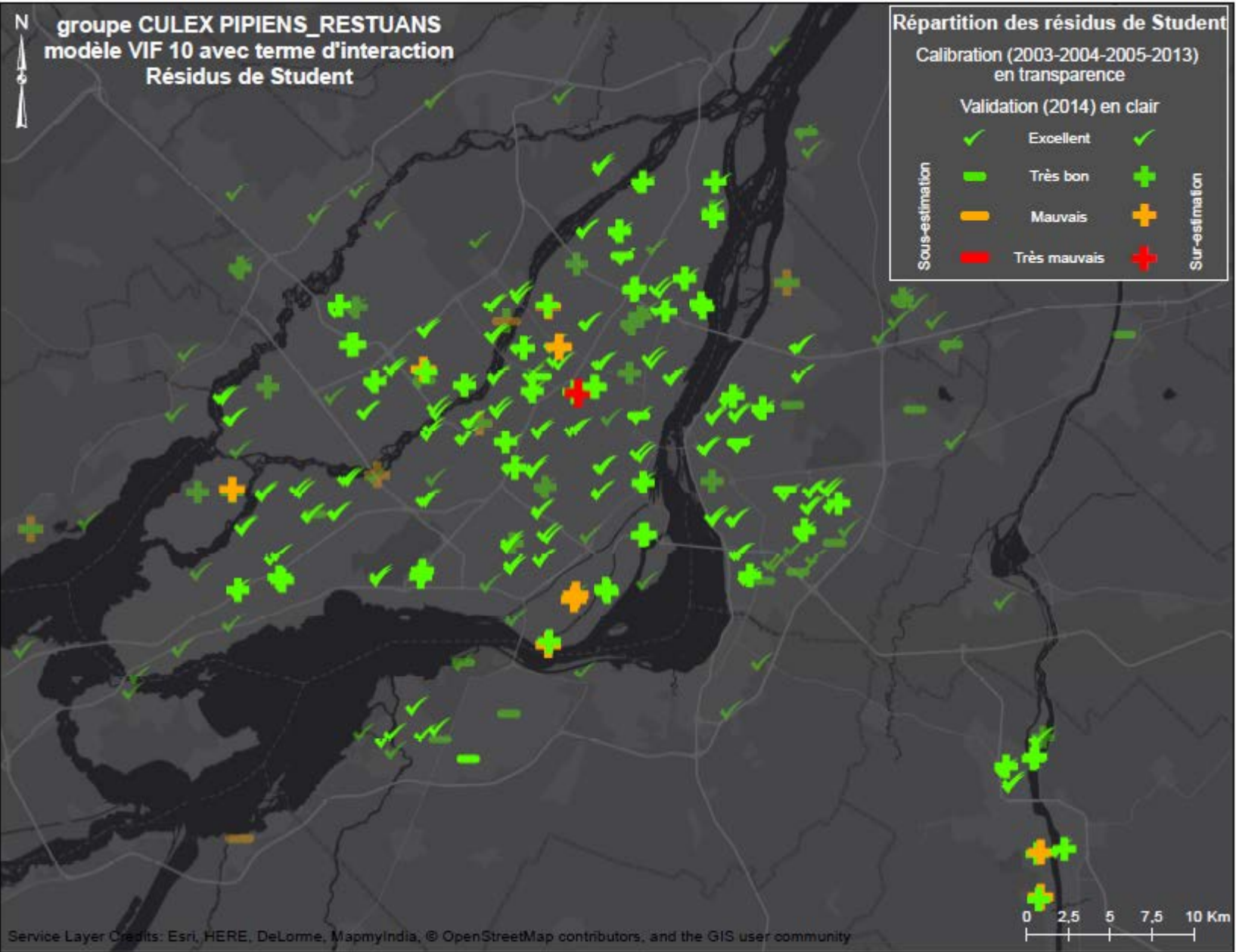
Pièges	Valeurs observées Box Cox	Valeurs prédites Box Cox	Résidus bruts	Résidus de Student
CHA001_0	10,87	6,06	-4,81	-0,95
CHA002_0	13,22	3,82	-9,40	-1,86
CHA003_0	11,53	6,81	-4,72	-0,93
CHA006_0	14,10	7,90	-6,20	-1,23
CHA007_0	14,60	5,01	-9,59	-1,90
CHA008_0	11,32	5,99	-5,33	-1,06
<b>Moyennes des 6 pièges CHA</b>	12,61	5,93	-6,67	-1,32
<b>Moyennes des 206 pièges</b>	<b>5,97</b>	<b>2,04</b>	<b>-3,93</b>	<b>-0,78</b>

Tableau 37 : Comparaison des valeurs du MLMIf d'*Aedes vexans* pour les pièges du secteur de la réserve faunique de Marguerite d'Youville par rapport à la moyenne de tous les pièges.

Pièges	Valeurs observées Box Cox	Valeurs prédites Box Cox	Résidus bruts	Résidus de Student
CHA001_0	5,14	3,59	-1,55	-2,79
CHA002_0	4,30	3,56	-0,74	-1,33
CHA003_0	4,75	3,58	-1,17	-2,11
CHA006_0	4,46	3,60	-0,86	-1,54
CHA007_0	4,56	3,56	-0,99	-1,79
CHA008_0	4,37	3,57	-0,79	-1,42
<b>Moyennes des 6 pièges CHA</b>	4,60	3,58	-1,02	-1,83
<b>Moyennes des 206 pièges</b>	<b>3,90</b>	<b>3,58</b>	<b>-0,32</b>	<b>-0,57</b>



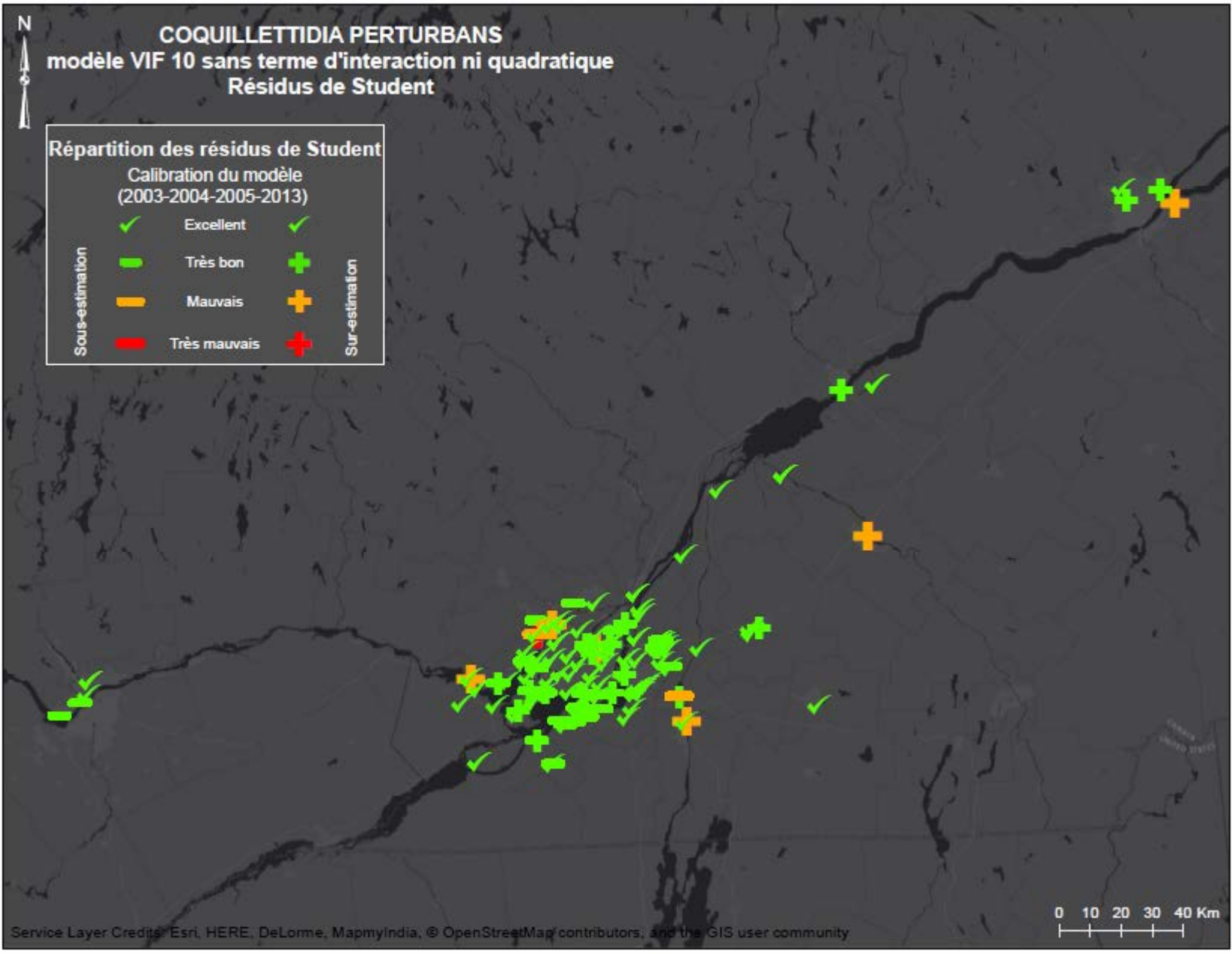
A



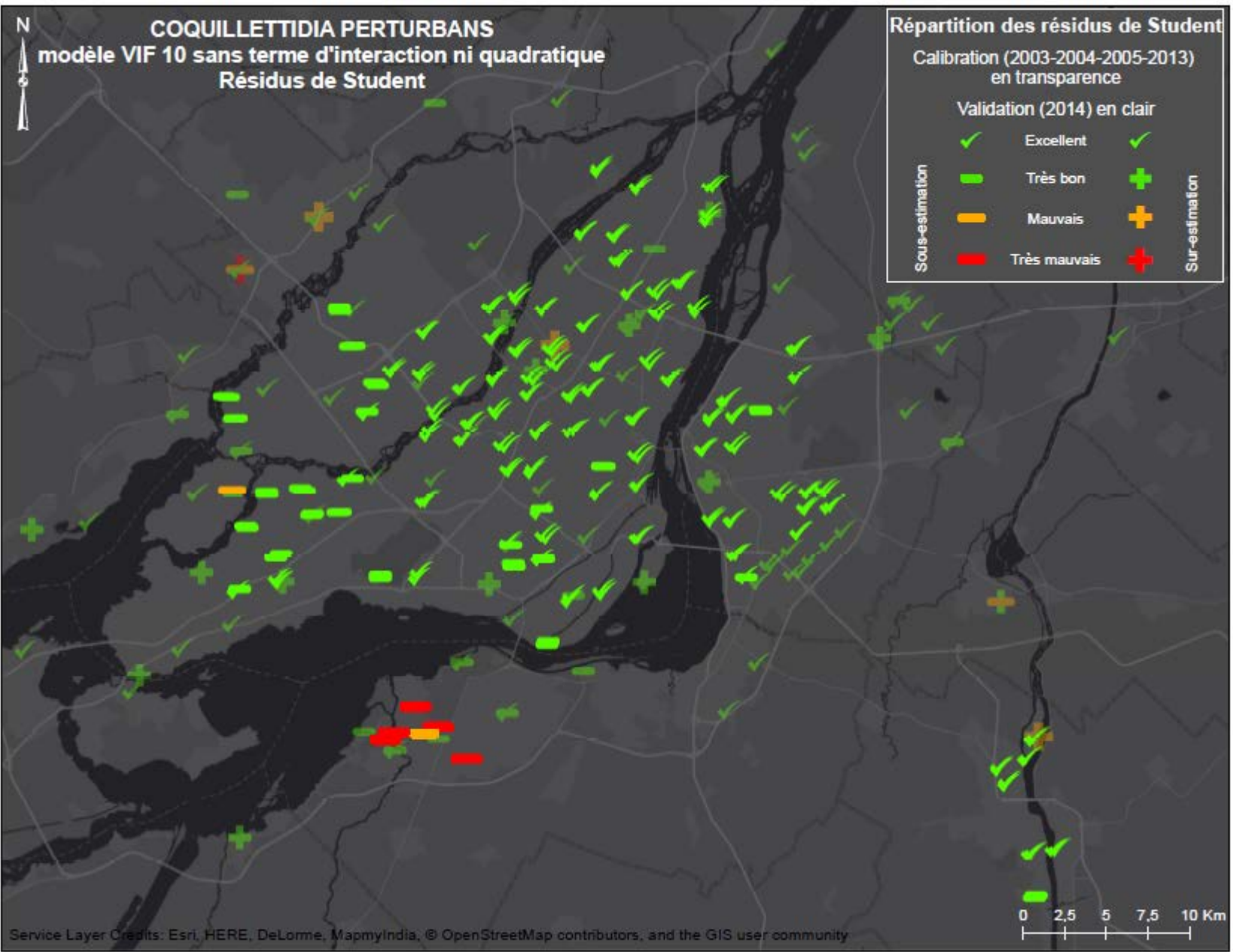
B

Figures 28A et B : Distribution spatiale des résidus de Student du MLMIf du groupe Culex pipiens-restuans pour les données d'entraînement (A) et superposant les données d'entraînement et de validation (B).





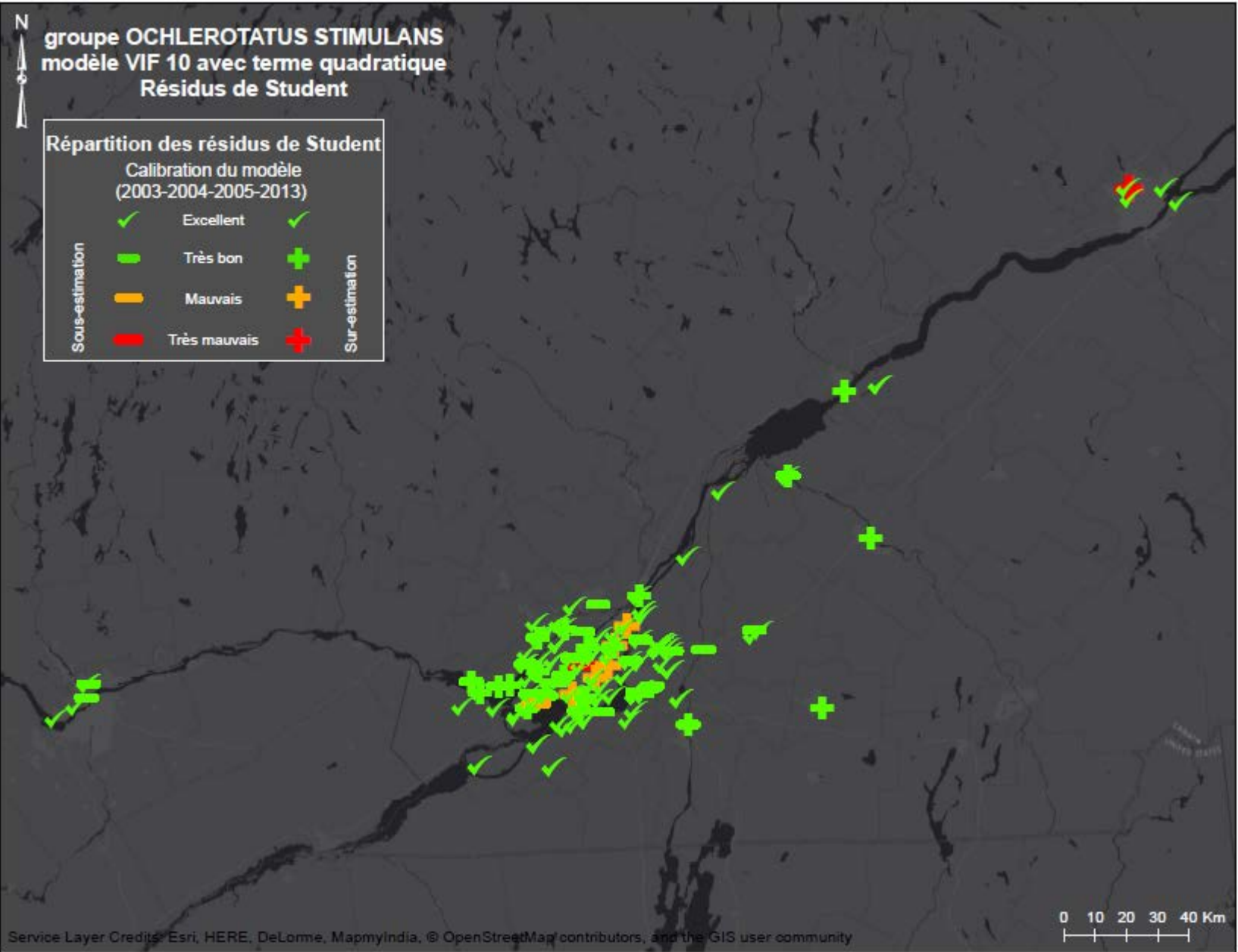
A



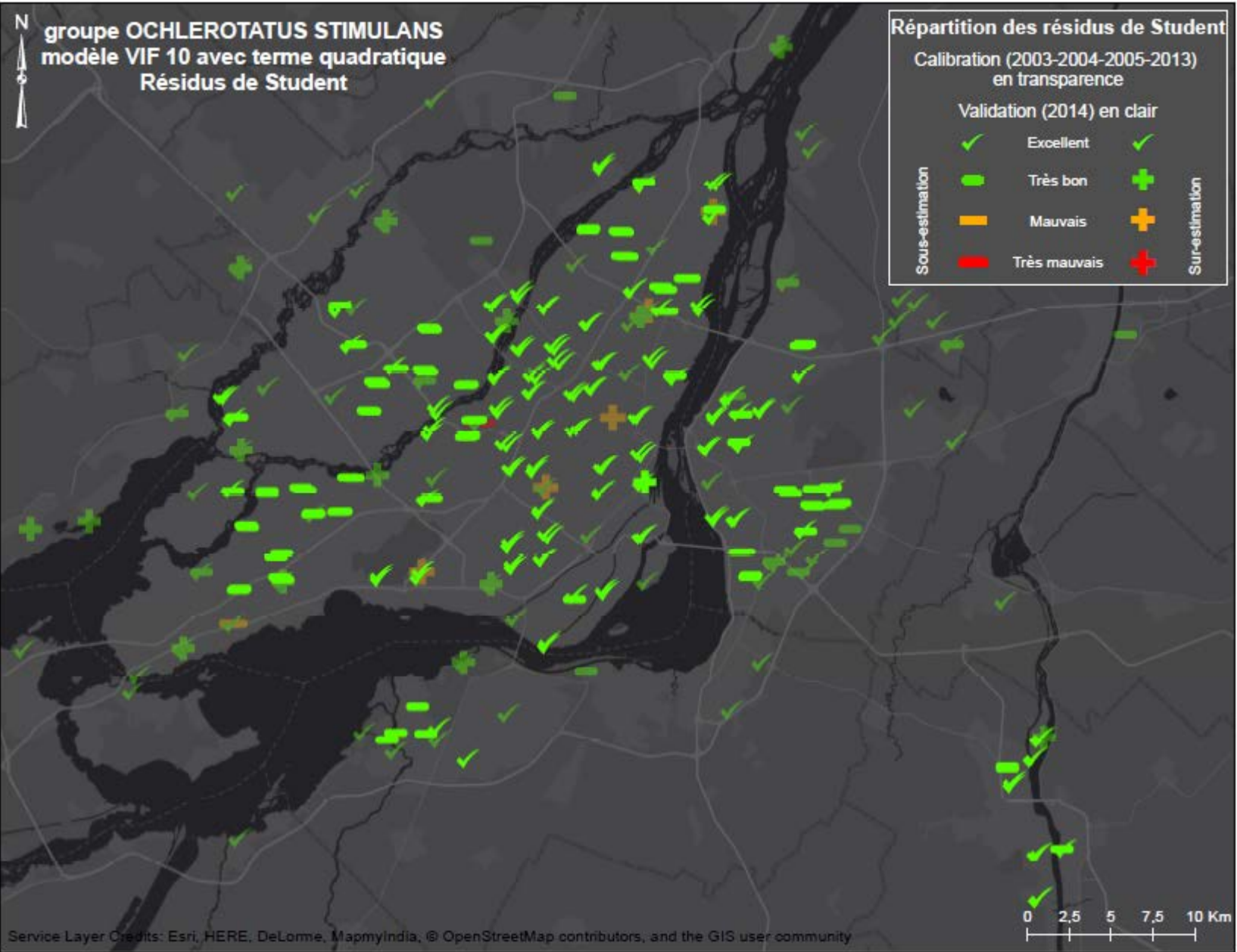
B

Figures 29A et B : Distribution spatiale des résidus de Student du MLM1f de *Coquillettidia perturbans* pour les données d'entraînement (A) et superposant les données d'entraînement et de validation (B).





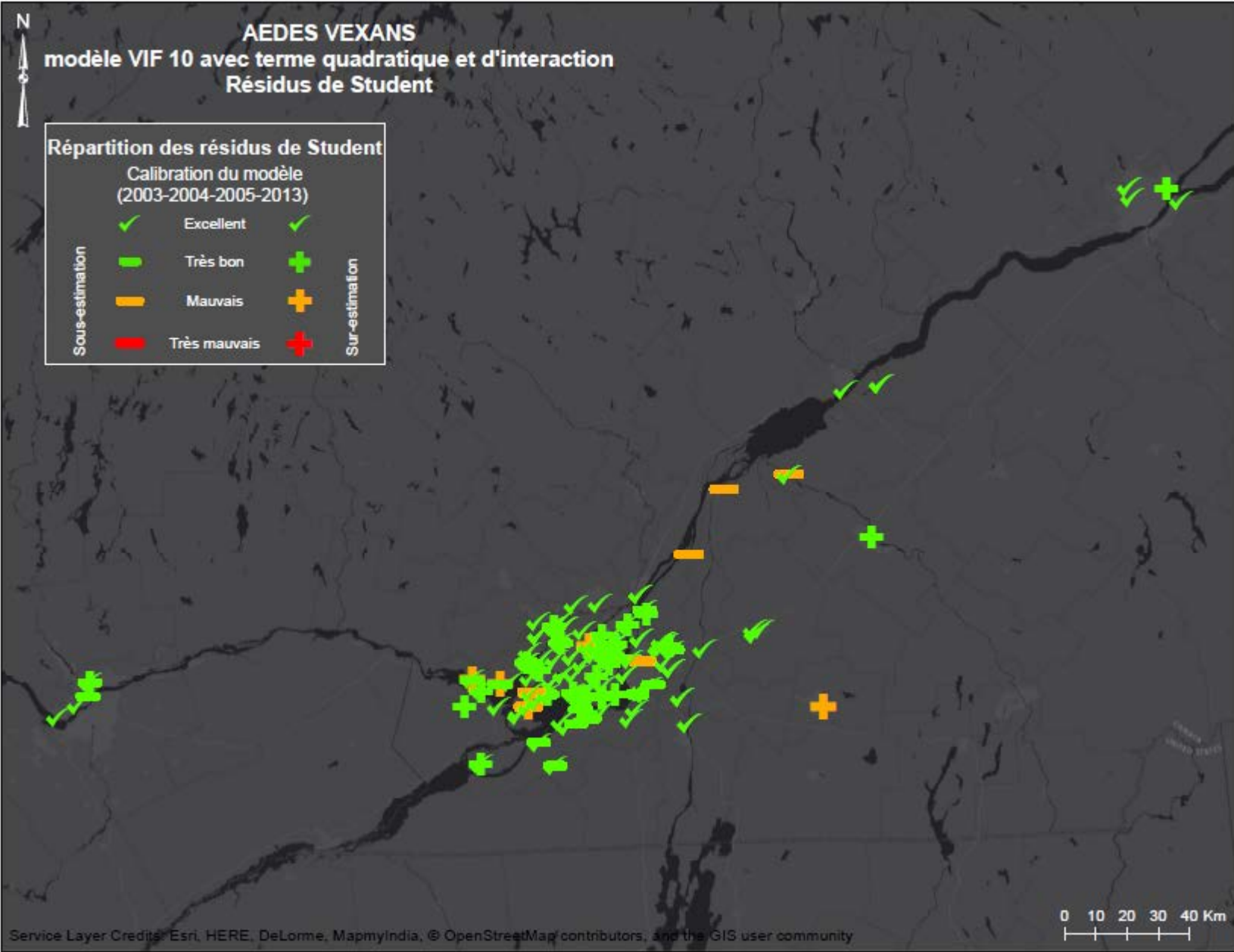
A



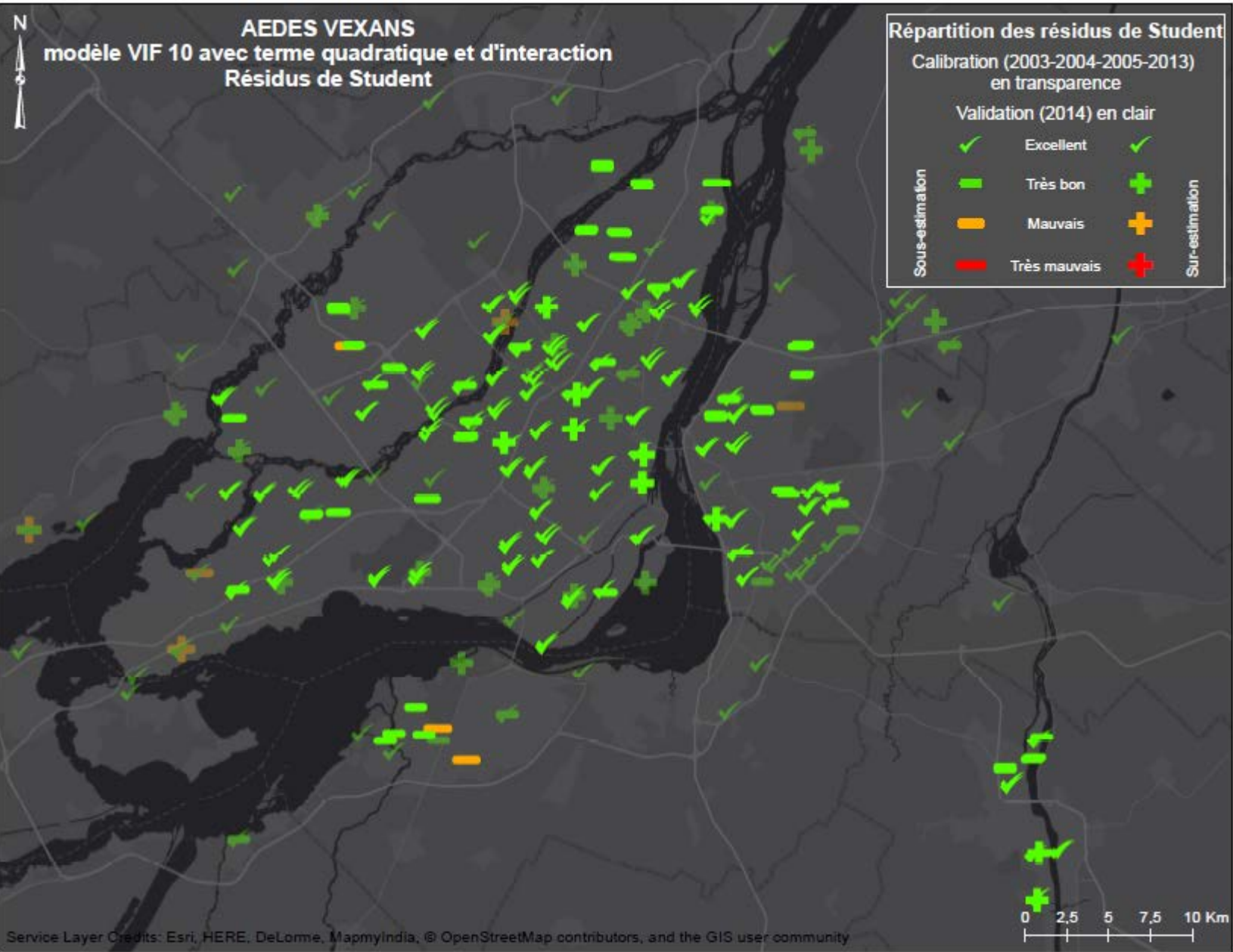
B

Figures 30A et B : Distribution spatiale des résidus de Student du MLMIf du groupe Ochlerotatus stimulans pour les données d'entraînement (A) et superposant les données d'entraînement et de validation (B).





A



B

Figure 31A et B : Distribution spatiale des résidus de Student du MLM1f d'Aedes vexans pour les données d'entraînement (A) et superposant les données d'entraînement et de validation (B).

### 5.5.2. Étape 2 : cartes annuelles des densités d'espèces dans la zone d'étude

Toutes les images de COS entrant comme variable dans les quatre MLM1f sont traitées selon le processus illustré par les figures 32A, B et C. L'effet de flou des figures 32B et 32C est causé par l'effet de lissage provoqué par le moyennage proximal injecté dans chaque pixel. La différence de coloration entre les deux dernières images est due à la différence d'échelle des valeurs. Pour l'exemple, la figure 32C, ici agrandie sur une portion de l'île de Montréal, montre le résultat pour la COS urbain qui entre dans le MLM1f de CQP et STMg. Le tableau 38 présente les moyennes et les écart types des variables simples d'entraînement appliquées aux images de validation correspondantes et le tableau 39 les paramètres pour les variables quadratiques et les termes d'interaction. Les valeurs parfois très élevées des écart-types reflètent la forte variabilité du paysage au sein de la zone d'étude. De plus, les écart-types plus grands pour les COS que pour les données météorologiques s'expliquent sans doute par la grande différence de précision des images utilisées en entrée.

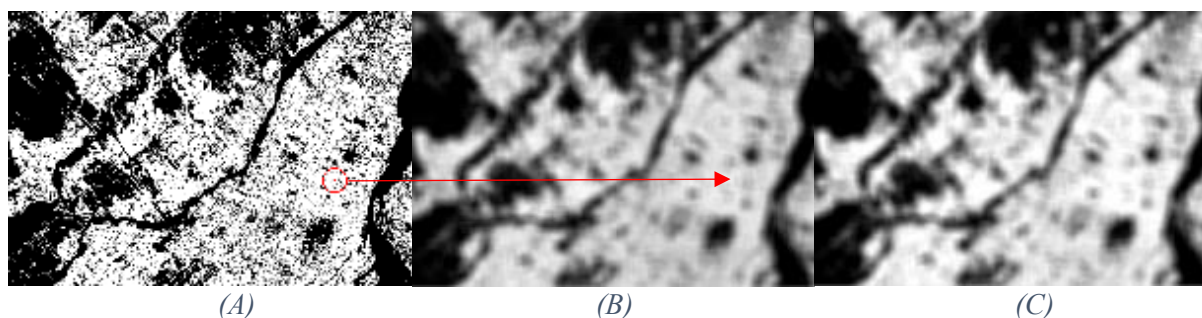


Figure 32A, B et C : Image d'origine de la classe Urbain (valeur en pourcentage d'occupation du sol de chaque pixel : 0 ou 100) (A), image moyennée de la classe Urbain (valeur en pourcentage de chaque pixel : [0 ; 100]) (B) et image centrée réduite de la classe Urbain (valeur de chaque pixel : [-2,34 ; 2,18]) (C).

Tableau 38 : Moyennes et écart-types des variables simple d'entraînement appliqués pour centrer réduire les variables de validation correspondantes.

Variable	Urbain	Agri	Prairie	Boise	TNsep_1	Pjan	Pfev	Pjun	Paou	Psep
Moyenne	51,77	6,26	3,72	15,00	11,24	42,67	51,15	94,84	104,76	106,18
Écart-type	22,08	11,79	5,15	14,95	1,06	15,48	11,77	33,84	30,17	25,87

Tableau 39 : Moyennes et écart-types des variables composées d'entraînement appliqués pour centrer réduire les variables de validation correspondantes.

Variable	Agri <sup>2</sup>	Boise <sup>2</sup>	Boise x TNsep_1	Pjan x Psep
Moyenne	177,33	447,43	167,74	4604,21
Écart-type	559,59	845,23	163,88	2064,87

En prenant le cas de CPRg comme exemple, la figure 33 illustre l'application mathématique de l'équation 1 du MLM1f correspondant aux images de COS préparées selon le procédé illustré par les figures 32A à C ainsi qu'aux images météorologiques que le modèle requière. Le résultat de ce calcul matriciel est la carte annuelle des abondances Box Cox de CPRg (figure 34). Cette carte est donc construite en multipliant les coefficients de participation correspondants aux images d'Agri 2014, de Prairie 2014, de Boise 2014, de TNsep\_1, soit 2013, ainsi que du terme d'interaction Boise x TNsep\_1.

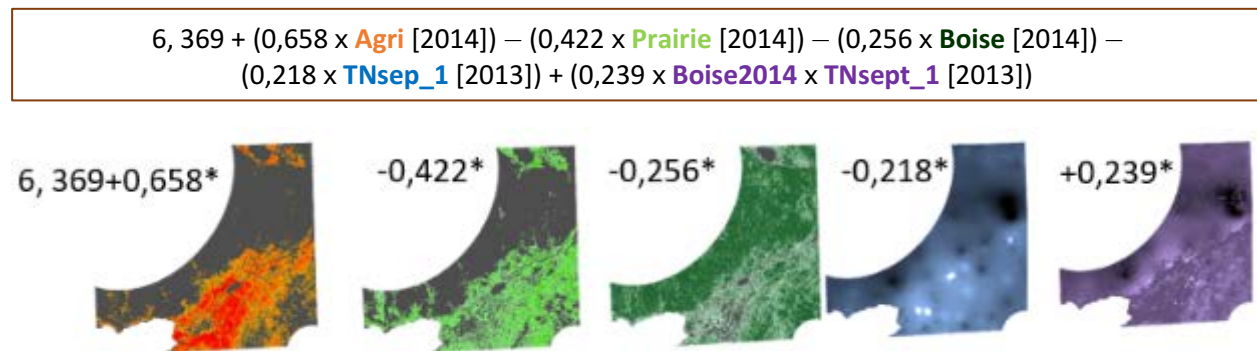


Figure 33 : Application du modèle final (MLM1f) du groupe *Culex pipiens-restuans* aux données environnementales de 2014.

Les cartes résultant de l'application des équations des MLM1f de chaque espèce [figures 34 à 37 (A et B)] prédisent, pour 2014, où se situent les sites les plus à risque de voir éclore une abondante population de moustiques de l'espèce en question. Afin de concentrer la lecture sur les milieux urbains et périurbains, d'où les pièges sont majoritairement issus, nous avons appliqué un masque blanchâtre en dehors de ces zones. En effet, la précision des prédictions sous les masques n'est pas garantie, faute de données de validation et parce que les valeurs centrées réduites s'éloignent rapidement des conditions d'entraînement à mesure que l'on s'éloigne des zones urbaines. À cause de ces valeurs centrées réduites, il est également difficile d'interpréter simplement et clairement la participation de chacune des variables d'un modèle.

Les observations de 2014 ont été ajoutées sur la carte à titre indicatif, car les gradients de couleur entre les points et les images ne sont pas toujours parfaitement synchronisés à cause de la différence dans l'étendue de leurs plages de valeurs. De plus, nous avons d'un côté une image de valeurs prédites et de l'autre les valeurs observées et non plus des valeurs de résidus comme dans la phase de validation. Cependant, les *hot spots* des cartes devraient en principe recouper les plus fortes densités de moustiques observées.



- *Groupe Culex pipiens-restuans*

La cartographie prédictive de 2014 présente une certaine homogénéité sur l'île de Montréal (en jaune) avec quelques *hot spots*, symbolisés en rouge, très localisés (figures 34A et B). En dehors de l'île, on observe un peu plus de variabilité. Les *hots spots* sur la zone d'étude se situent principalement dans les zones à forte densité de boisés. Or le coefficient de participation de cette variable est négatif dans le MLM1f de CPRg (équation 1). Cela signifie que c'est le terme d'interaction, composé par les boisés et les TNsep\_1 qui prend le dessus sur la variable simple boisé. La couleur orangée intermédiaire est principalement expliquée par la présence de parcelles agricoles qui a une forte influence positive dans le modèle, ce qui semble donc se confirmer visuellement. Les quelques rares *cold spots* symbolisés en vert sont alimentés par une couverture dense de prairies qui ont une influence négative dans le modèle. Dans les Laurentides, le *cold spot* est expliqué par les valeurs de températures drastiquement plus faibles à cet endroit ainsi que par la présence accrue de boisés, ce qui fait du terme d'interaction une variable particulièrement défavorable aux CPRg à cet endroit.

- *Coquillettidia perturbans*

L'interprétation de la carte de CQP (figures 35A et B) est relativement simple, puisqu'il apparaît assez clairement que les abondances de moustiques sont inversement proportionnelles à la densité urbaine. Cette composante agit négativement dans les résultats du MLM1f de CQP (équation 2). Ainsi, même les routes, de valeur plus faible, semblent se découper dans l'uniformité du paysage non urbain. Comme cette COS est la seule qui intervient dans le modèle, les variations d'intensité plus subtiles qui se devinent d'un milieu urbain à un autre sont influencées par les variations des deux variables de précipitations. La couleur des points d'observation se fond ici dans celles de la carte, ce qui démontre la précision de la cartographie.

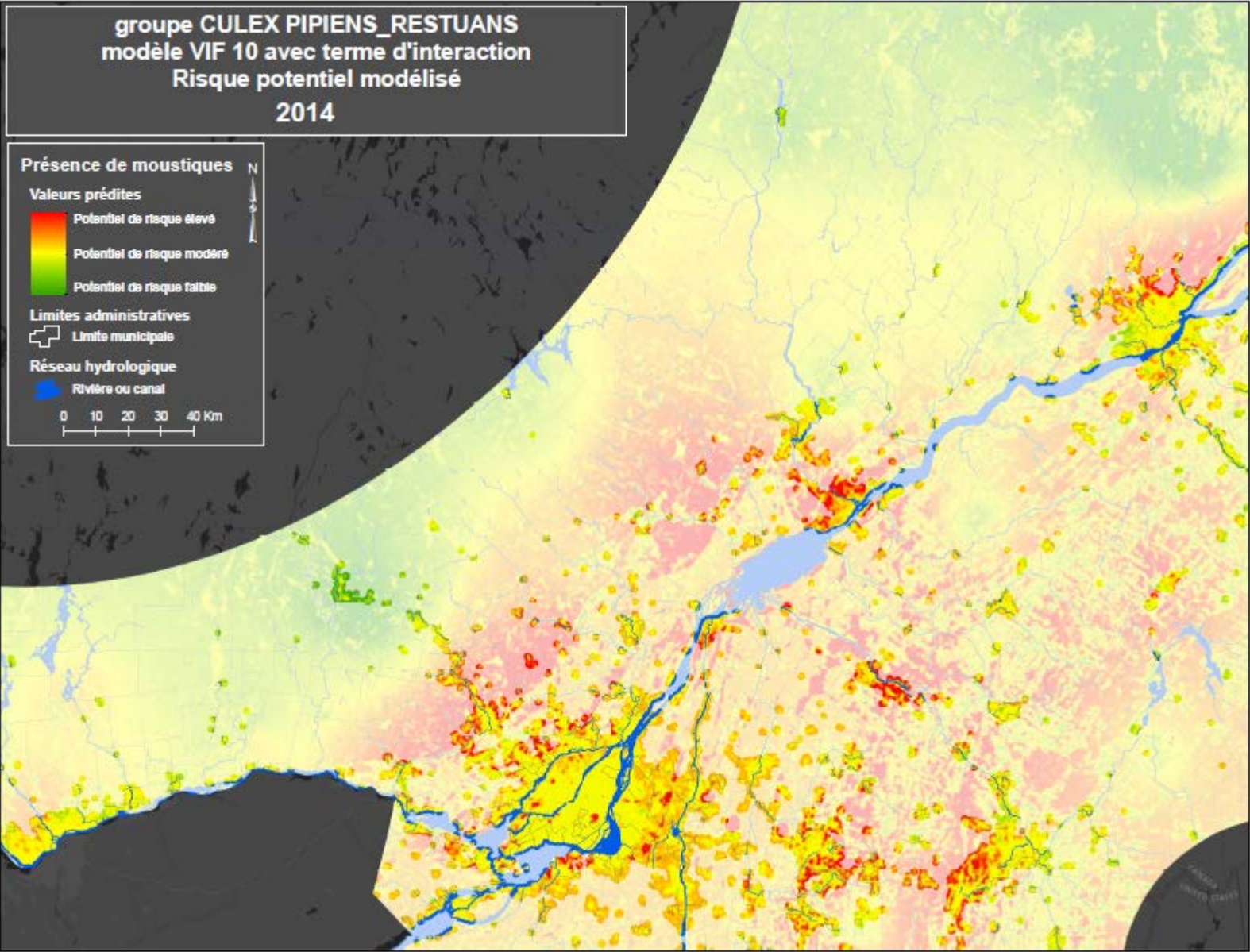
- *Groupe Ochlerotatus stimulans*

Tout comme pour les CQP, l'influence négative des milieux urbains sur les abondances de STMg est flagrante sur les figures 36A et B. L'influence des boisés y est également appréciable. Les coefficients de participation d'influence contraire entre la variable simple et son terme quadratique créent une sorte d'effet de bordure en lisière des boisés en dessinant des variations très rapides de valeurs qui prennent la forme de lignes rouges périphériques plus ou moins

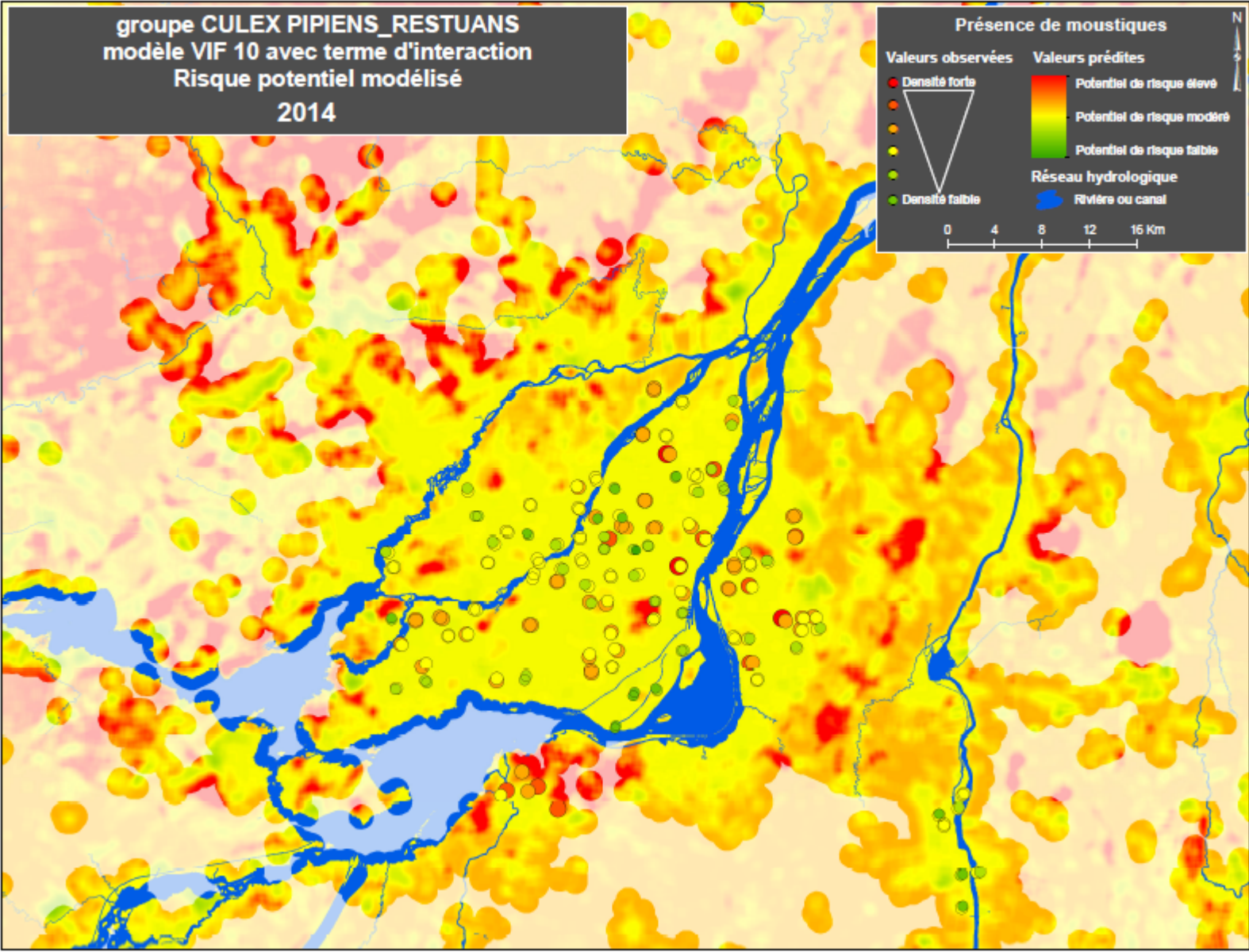
épaisses. Cet effet est plus visible sous le masque, loin des milieux urbains. La participation des précipitations de mois de juin est plus subtile. Cela est dû en grande partie parce que les variables météorologiques sont continues et que la variation plus fluide de leurs valeurs est moins remarquable que les différences plus drastiques de changement de COS. Dans le MLM1f de CQP, les précipitations sont en outre les variables qui ont le plus faible coefficient de participation, ce qui rend leur influence moins facilement perceptible pour cette raison également.

- *Aedes vexans*

La variabilité des abondances de VEX sur le territoire présente une variabilité très rapide (en rouge) à la lisière des zones agricoles (figures 37A et B). Comme pour STMg avec la variable boisé, cet effet est causé par les influences contraires de la variable de COS agricole, dont le coefficient de participation est positif, et de son terme quadratique dont les valeurs varient plus rapidement, mais dont le coefficient de participation est plus faible. Comme c'est la seule COS qui intervient dans le MLM1f de VEX, cet effet de bordure n'est pas atténué comme pour STMg. Ainsi, les lignes rouges semblent plus nettes autour des zones agricoles pour VEX qu'autour des boisés pour STMg. Ici aussi l'influence des températures est beaucoup plus subtile, mais explique sans doute les différences d'intensité dans les gammes de couleurs entre des zones apparemment semblables. On remarque une telle nuance entre le nord du fleuve aux alentours de La Présentation et le sud du fleuve un peu au sud de La Prairie.



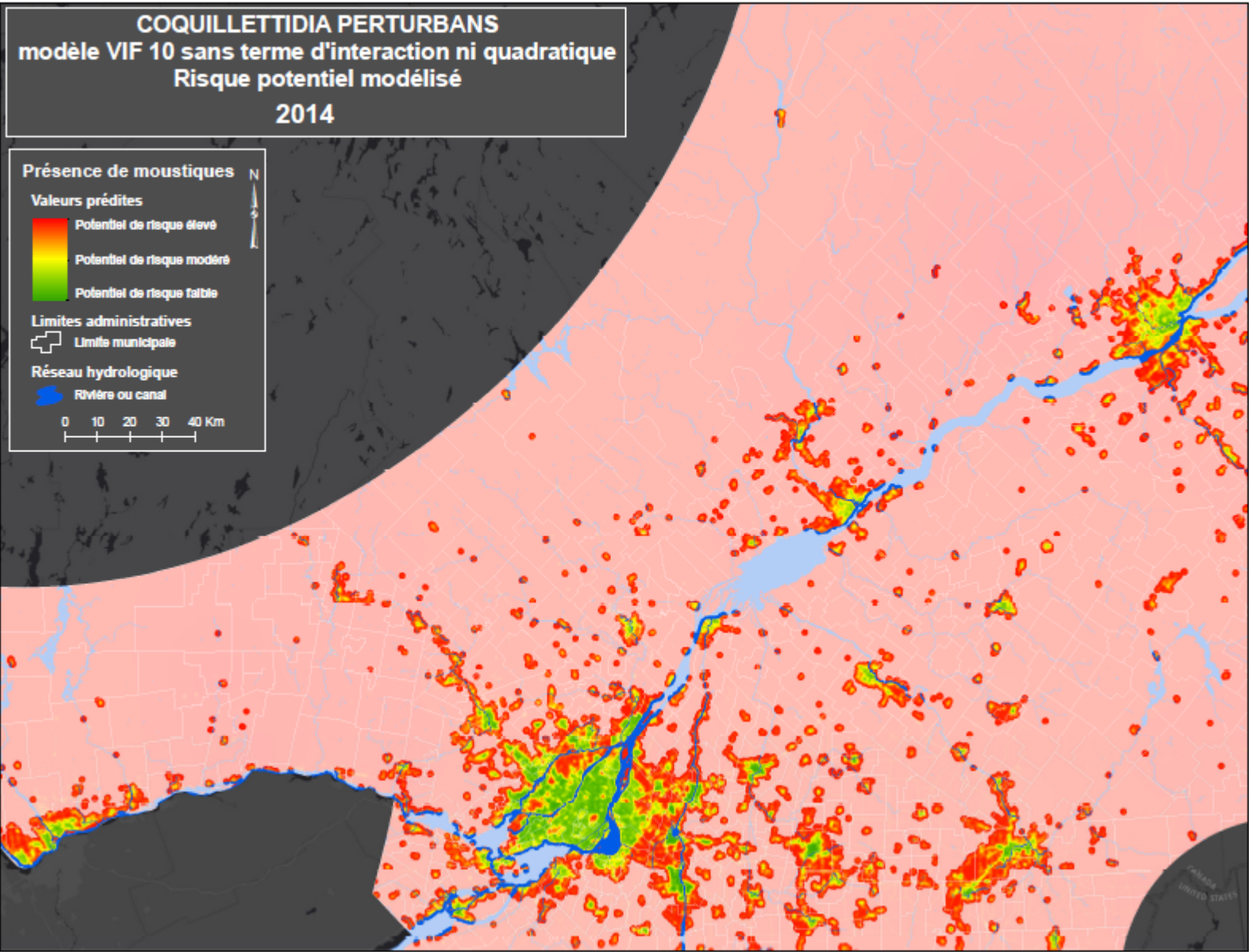
A



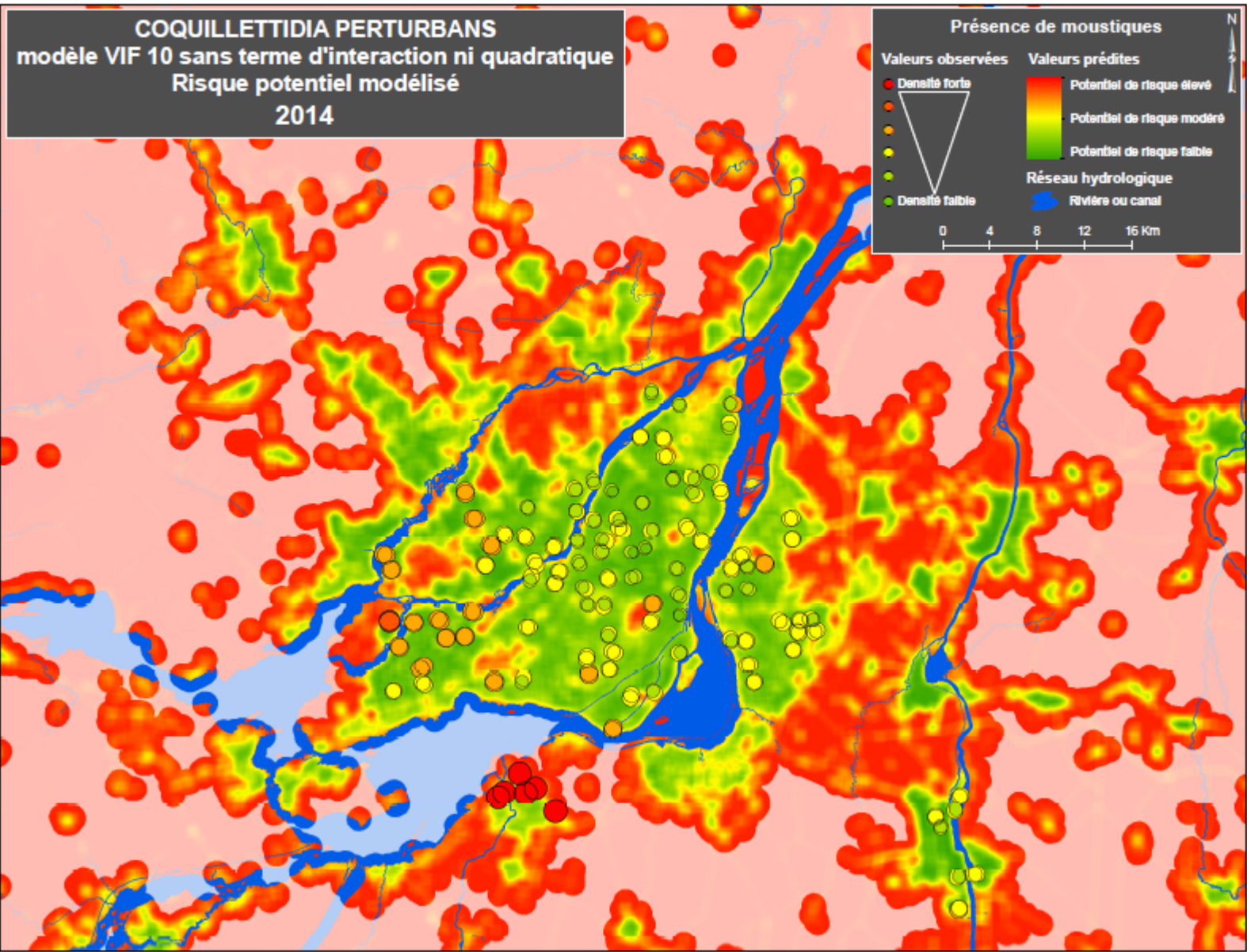
B

Figures 34A et B : Cartographie annuelle prédictive du groupe Culex pipiens-restuans validée avec les données de 2014 sur la zone d'étude (A) et agrandie sur le grand Montréal (B).





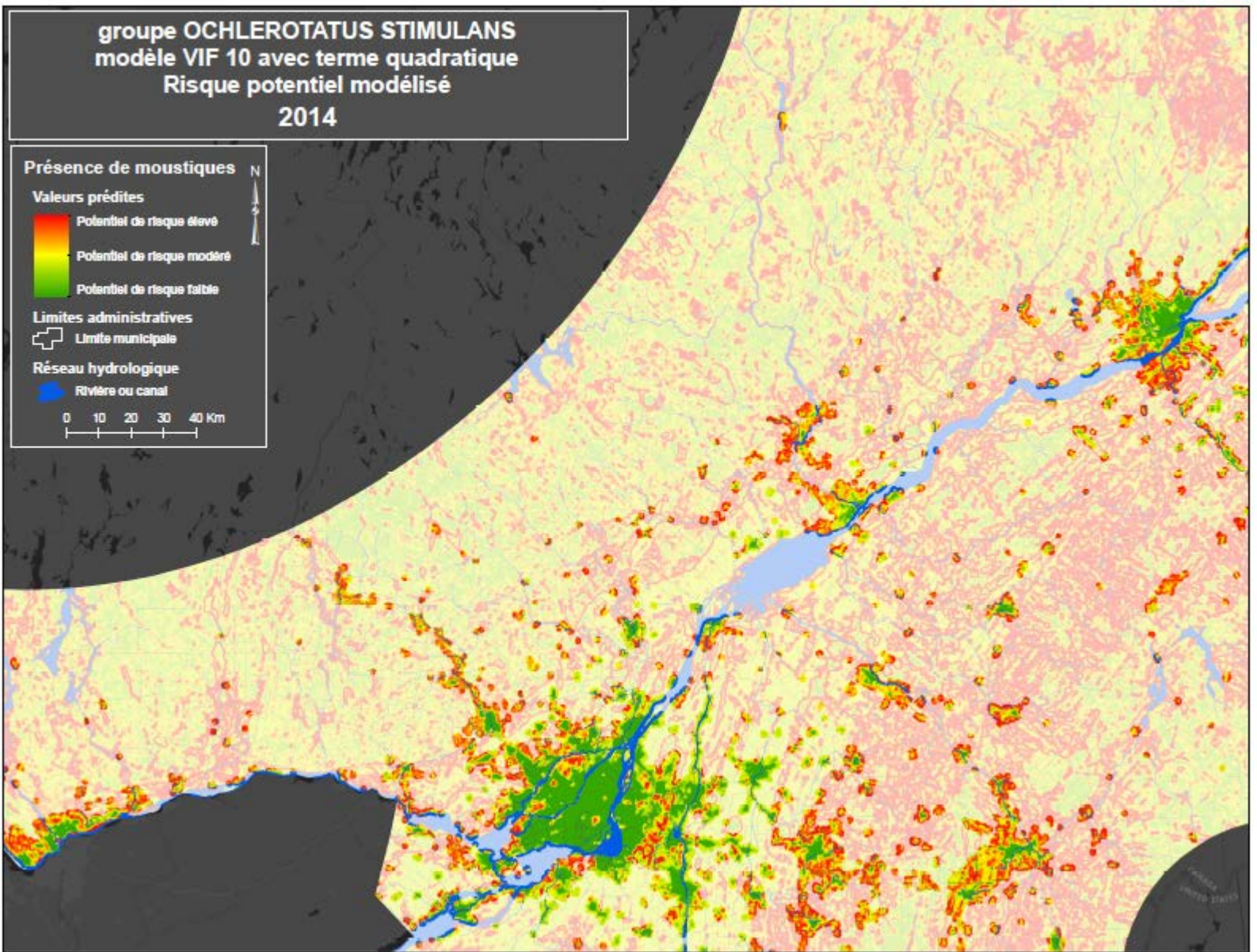
A



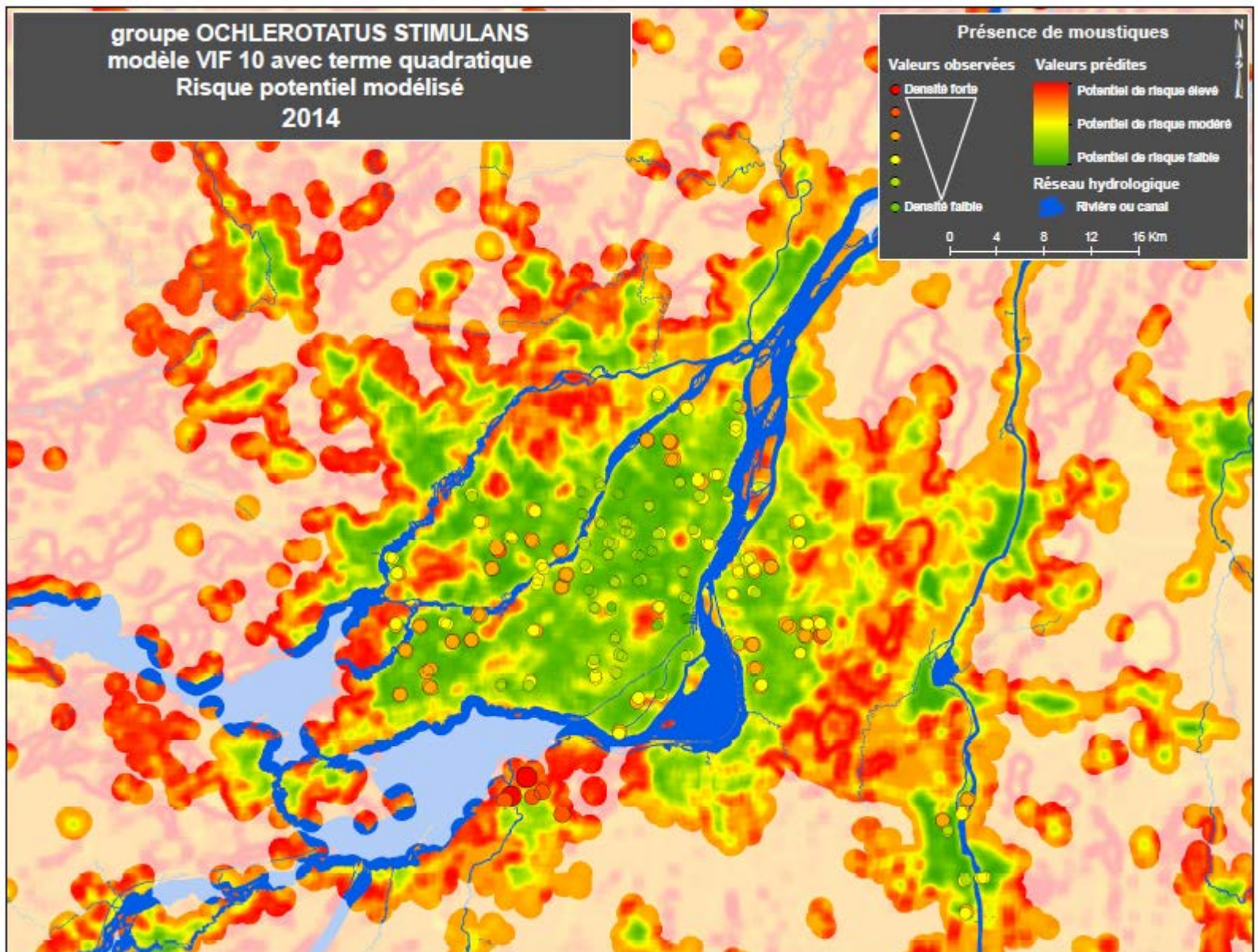
B

Figures 35A et B : Cartographie annuelle prédictive de Coquillettidia perturbans validée avec les données de 2014 sur la zone d'étude (A) et agrandie sur le grand Montréal (B).





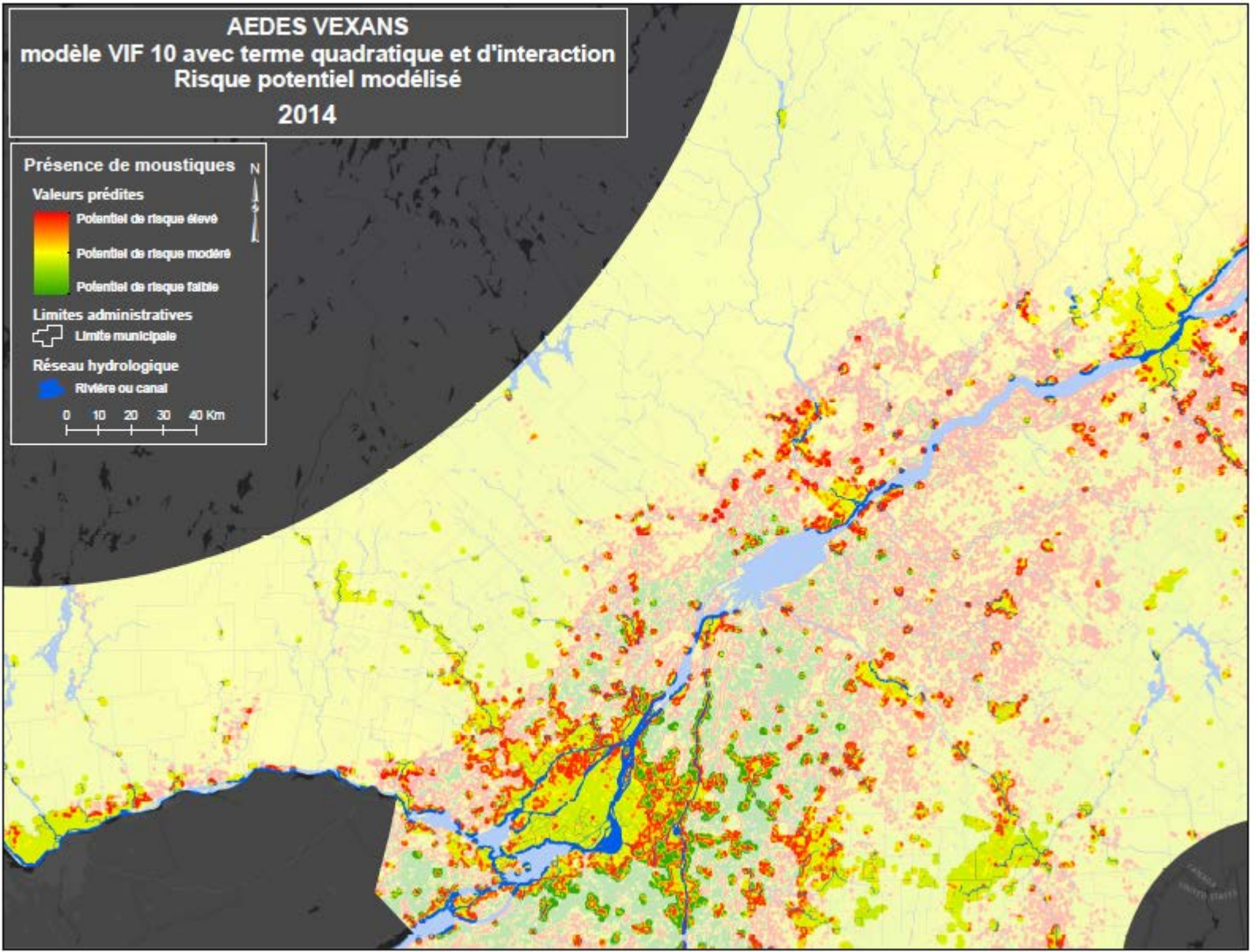
A



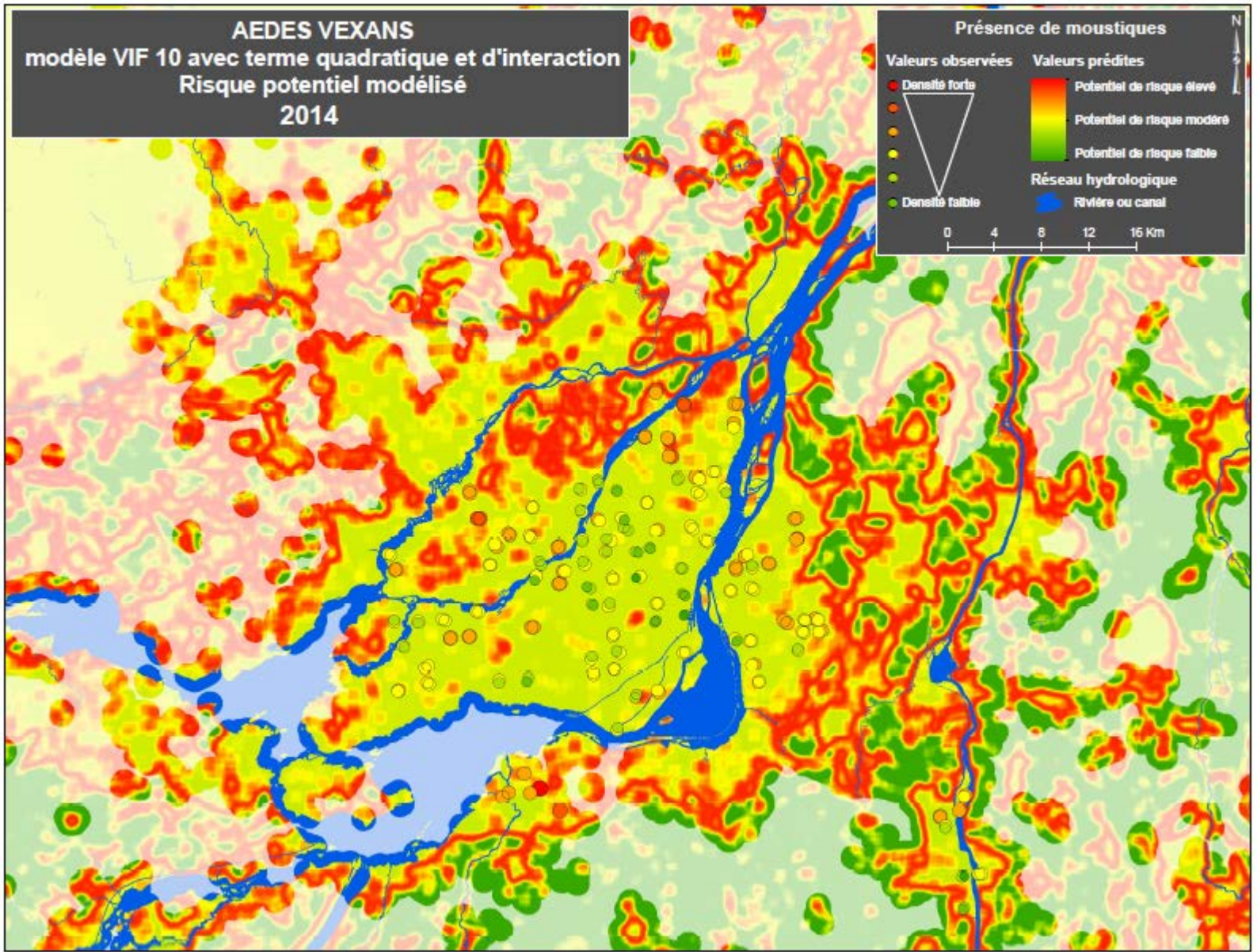
B

Figures 36A et B : Cartographie annuelle prédictive du groupe Ochlerotatus stimulans validée avec les données de 2014 sur la zone d'étude (A) et agrandie sur le grand Montréal (B).





A



B

Figures 37A et B : Cartographie annuelle prédictive d'Aedes vexans validée avec les données de 2014 sur la zone d'étude (A) et agrandie sur le grand Montréal (B).



## 6. Discussion

Notre projet a abouti à la production de cartes de risques calculées à l'aide d'un modèle de régression linéaire qui s'est révélé robuste pour chacune des quatre espèces étudiées, soit CPRg, CQP, STMg et VEX. Dans les quatre cas, nous avons trouvé que les modèles dont les variables ont été sélectionnées au seuil VIF égal à 10 sont plus performants que ceux dont les variables ont été sélectionnées aux seuils VIF 5 ou 3. Non seulement cela démontre que les modèles incluant des variables météorologiques associées à des COS (VIF 10 et VIF 5) sont meilleurs que ceux basés uniquement sur les COS (VIF 3), mais ce résultat indique en plus que l'utilisation d'un seuil VIF trop restrictif est susceptible de nuire à la qualité de la modélisation en écartant des variables qui ont un rôle statistiquement significatif.

Les étapes de validation interne et externe ont donné des résultats satisfaisants en termes de distribution des résidus de Student. Cette distribution a ensuite été validée à la fois sur le plan statistique et sur le plan spatial. En outre, l'utilisation des données de 2014, choisies pour la validation comparativement à des années antérieures choisies pour l'entraînement des modèles, a également permis de valider l'aspect temporel, donc prédictif de chacun des quatre MLM1f. Ainsi, les cartes finales de risque reposent sur une méthode logique qui a démontré son efficacité pour quatre espèces de moustiques très différentes.

La cartographie prédictive des abondances de CPRg se base sur les variables de COS agricole, de prairie et de boisé ainsi que sur les températures minimales moyennes du mois de septembre de l'année précédente et de l'interaction entre ces températures et les boisés. Bien que ce groupe de moustiques soit réputé pour être fortement présent en ville (Jacob *et al.*, 2009b; Harris *et al.*, 2015) démontrent toutefois que la distribution de cette espèce est très fortement influencée par la composition de l'environnement au sein de son habitat partiellement urbain. De plus, l'étude de Gardner *et al.* (2014) en Illinois, avance que le groupe CPRg est présent en plus forte proportion dans les milieux boisés que dans les milieux agricoles, les prairies ou les milieux urbains qu'il affectionne également. Cette étude démontre également que ces quatre COS ont la même importance pour les *Culex pipiens* et les *Culex restuans* testés individuellement. Notre étude démontre toutefois que, pour le sud du Québec, ce sont les variations de couverture du milieu agricole qui ont le plus d'influence sur la disparité de leur abondance. Différents modèles de prédiction des abondances de CPRg n'utilisent pas de COS, mais des variables météorologiques

basées sur les intervalles de temps entre précipitations et pics d'abondance (Hongoh *et al.*, 2012; Lebl *et al.*, 2013; Ripoché *et al.*, 2019). Trawinski and Mackay (2010) utilisent l'occupation du sol en plus des conditions météorologiques. Leurs résultats indiquent que CPRg est corrélé positivement avec les milieux urbains purs ou mixtes, les milieux agricoles et les zones d'herbacées et négativement avec les couvertures forestières, ce qui conforte nos résultats concernant l'agriculture et les boisés.

Les études concernant les habitats de VEX indiquent que cette espèce se développe principalement dans des cuvettes d'eau temporaires causées par des précipitations ou de récentes inondations (Crans *et al.*, 2014; Trawinski and Mackay, 2009). Ces mares et autres flaques peuvent être situées dans les milieux agricoles (Rocheleau, 2016), les boisées, les prairies (Gardner *et al.*, 2014; Strickman, 1980 et 1982) ou les pâturages (Strickman, 1980 et 1982). Cela concorde en partie avec nos résultats qui ont révélé l'influence du milieu agricole conjointement avec les précipitations hivernales et automnales. Il existe peu de modélisation appliquée à l'espèce VEX pour le Canada. Ripoché *et al.* (2019) ont basé leur étude sur les conditions météorologiques et l'utilisation de larvicides pour les CPRg et VEX en Ontario, mais leurs modèles ne considèrent pas de variables liées à l'occupation du sol.

Quant à l'espèce STMg, nous avons trouvé qu'elle est plutôt influencée par les variations au sein des zones en partie boisées et urbaines associées négativement avec les précipitations du mois de juin. Ce résultat ne recoupe pas les rares informations empiriques de la littérature qui le décrivent comme recherchant les mares printanières (Crans *et al.*, 2004; Maire et Aubin, 1980; Wood *et al.*, 1979), bien que Crans *et al.* (2014) situe l'habitat larvaire d'*Ochlerotatus stimulans* dans les dépressions des milieux boisés. Nous ne possédons toutefois pas d'autres informations ni de résultats provenant d'autres essais de modélisation pour cette espèce. Nos données de capture regroupent les *Ochlerotatus stimulans* et les *hexodontus* dans le groupe STMg. Il semblerait que les *Ochlerotatus hexodontus* se développent dans les zones herbacées inondées (Albers and Bradley, 2011). Or, l'étude de Ward and Darsie (2005) situe la zone de distribution des HEX plus au nord du Québec. Ainsi, il est possible que nos échantillons de STMg ne contiennent pas ou peu de cette espèce.

Il est encore plus délicat de confronter nos résultats avec d'autres études concernant les CQP, car cette espèce est très peu documentée. L'article de Crans *et al.* (2004) mentionne que les œufs de CQP sont déposés directement sur l'eau et celui de Gardner *et al.* (2014) dénombre des CQP



capturés à l'aide de pièges LT. Ces moustiques ont été piégés principalement dans des prairies, puis dans des milieux boisés et des terres agricoles et, dans une moindre mesure, dans des quartiers résidentiels. Toutefois, cette étude a été conduite dans un comté de l'Illinois. Ce qui fait de notre projet, au regard de la qualité des modèles, un précieux outil de renseignement quant aux conditions environnementales de prédilection des CQP au Québec. Nous avons démontré que ces moustiques se développent surtout en périphérie des milieux urbains en lien avec les précipitations des mois de janvier et d'août.

Un des points saillants de ce travail de modélisation concerne la validation spatiale. En dehors de la zone autour du refuge de Châteauguay, nos cartes présentent une capacité prédictive qui a été validée sur toute la zone d'étude. La répartition des données d'entraînement sur le territoire allait de Gatineau jusqu'à Québec, ce qui est source d'une grande fluctuation dans les valeurs des variables, en particulier dans les données météorologiques. Nos modèles ont correctement intégré ces variations spatiales qui ont été traduites de façon précise en abondances de moustiques au sein d'un territoire plus local et principalement concentré sur la grande région de Montréal. Bien que nous n'ayons pas de données pour valider cette hypothèse, la différence notable entre les étendues des données d'entraînement et celle des données de validation nous permet donc de supposer que nos modèles pourraient être appliqués à des zones urbaines et périurbaines dans d'autres régions du sud du Québec.

Nos cartes de risque finales mettent également en lumière une hétérogénéité marquée entre les milieux urbains et périurbains au niveau des abondances prédites, et ce, pour toutes les espèces, même celles, comme CPRg et VEX, dont la variable Urbain n'entre pas dans le modèle. Cela s'explique peut-être par le biais induit par la localisation des pièges majoritairement posés à proximité des zones habitées, donc en milieu urbain ou périurbain. En fait, pour les années d'entraînement, absolument tous nos pièges sont à moins de 1 km d'une zone urbaine. Le pourcentage de présence de cette COS au sein des zones tampons est au minimum de 1,5 %, alors que toutes les autres COS comportent des valeurs nulles. C'est pourquoi, dans le cadre d'une modélisation basée sur des COS, et pour affiner les résultats en milieu purement urbain, nous recommandons d'utiliser des données permettant de distinguer différents paliers de densité urbaine, à une résolution de 30 m ou moins. Cette suggestion est d'autant plus valable qu'il a déjà été démontré que cela a un impact sur certaines espèces (Marcantonio *et al.*, 2015). D'autre part, dans le but de mieux protéger les humains à risque, certaines études préconisent jusqu'à

l'intégration de paramètres socio-économiques (Kotchi *et al.*, 2016 ; Tachiiri *et al.*, 2006). Une autre approche serait de localiser les cas de VNO déclarés par année sur les cartes correspondantes. Cependant, le cadre de la transmission du VNO dépend de facteurs complexes et l'abondance des moustiques vecteurs de VNO ne signifie pas systématiquement une forte transmission du virus. Le risque est toutefois naturellement plus fort s'il y a plus de moustiques.

Nos données de COS n'étaient disponibles qu'à compter de 2011. Or, le paysage a évolué entre 2003 et 2011. Nous recommandons donc d'employer des sources de données de COS multitemporelles qui puissent offrir une classification fine du territoire et dont la résolution est de 30 m ou moins, car nous avons démontré l'importance d'utiliser des COS autant que les données de température et de précipitation dans une modélisation d'abondances de moustiques.

Étant donnée la nature de nos données météorologiques, et bien que nos résultats soient très encourageants, nous conseillons d'affiner la précision des variables météorologiques en utilisant des données microclimatiques. Ces deux aspects combinés permettraient ainsi éventuellement de capturer plus finement les variations d'abondances de moustiques pour cartographier les risques associés à l'échelle municipale par exemple. Aussi, de nouvelles variables météorologiques explicatives pourraient être testées, comme le vent (Hamer *et al.* 2014, Karki *et al.*, 2016, Lebl *et al.*, 2013) ou l'humidité relative (Platt *et al.*, 1958). Cela dit, nos modèles se sont montrés performants, même avec de simples variables génériques de températures et de précipitations qui ont su répondre aux caractéristiques propres à chacune des espèces.

Outre le fait d'avoir réussi à modéliser les abondances de quatre espèces différentes, nous avons mis en avant un travail sur des espèces rarement exploitées alors qu'elles sont potentiellement vectrices de plusieurs virus transmissibles aux humains. En effet, beaucoup d'études se concentrent sur CPRg (Edillo *et al.*, 2009; Jacob *et al.*, 2009a et b, Lebl *et al.*, 2013) voire sur CPRg et VEX (Ripoche *et al.*, 2019) pour leur potentiel respectif à transmettre le VNO. D'autres études se penchent plus largement sur les espèces à risque de transmettre le VNO, mais sans tenir compte de la transmission éventuelle d'autres virus (Trawinsky and Mackay, 2009 et 2010; Turell *et al.*, 2001 et 2005; Wang *et al.*, 2010; Wang *et al.*, 2014; Wimberly *et al.*, 2014). Ainsi les espèces étudiées sont toujours plus ou moins les mêmes. Au contraire, notre sélection d'espèces s'est justement basée sur leur potentiel à transmettre plusieurs virus tels que le CEV, le CVV, l'EEEV, le JCV, le LACV, le SLEV et le SSHV (tableau 1). C'est ainsi que nous avons travaillé sur quatre espèces, soit CPRg, CQP, STMg et VEX, dont deux espèces, CQP et STMg,

n'ont encore jamais été modélisées, sur un unique schéma méthodologique commun. Cela ouvre la voie à travailler sur des espèces parfois délaissées, mais qui pourraient avoir un rôle à jouer dans la santé publique en lien avec les zoonoses d'autant que notre méthode suit une démarche relativement simple et reproductible.

Dans la littérature, les discussions restent très souvent ouvertes pour l'amélioration de chacun des modèles rencontrés. Par exemple, certaines études recommandent la prise en compte des effets de l'épandage de larvicide sur la densité d'une population pour prédire les générations suivantes (Wang *et al.*, 2011; Ripoche *et al.*, 2019). Bien que ces effets soient controversés par d'autres études, en particulier pour CPRg, (Bodner *et al.*, 2018), nos résultats obtenus aux abords du refuge faunique de l'île Saint-Bernard, où les épandages de larvicides sont peut-être restreints, amènent à se poser la question de l'impact réel de ces campagnes d'éradication de moustiques en fonction des espèces. C'est une donnée supplémentaire qui pourrait être intégrée comme variable explicative dans la méthodologie présentée, au même titre que la présence des refuges fauniques. Pour l'histoire, le refuge en question n'est protégé à perpétuité que depuis 2010. Un an plus tard, le secteur nommé le « tertre », situé à la pointe ouest de l'île a été acquis par la ville de Châteauguay (Île Saint-Bernard, 2018). Cependant, rien n'indique si des aménagements particuliers ont été effectués ou non depuis cette acquisition, qui pourraient expliquer la quantité de moustiques plus élevée que la moyenne à cet endroit en 2014. La question de la cause exacte de la sous-estimation systématique de nos modèles en périphérie du refuge reste ouverte, mais la présence de ces 223 ha riches en espèces fauniques et floristiques offre une piste d'explication.

Afin d'aller plus loin, et dans une optique d'étude de l'impact des changements climatiques sur la transmission de maladies par les moustiques dans le sud du Québec, nous avons démontré que nos modèles remplissent leur rôle prédictif. Il est donc envisageable d'utiliser des prédictions climatiques à long terme, comme celles à 50 ou à 100 ans associées à des images anticipant l'évolution du paysage sur la même période pour vérifier s'il y aura dans le futur une dispersion de chacune des espèces modélisées sur le territoire québécois.

## 7. Conclusion

La méthodologie développée dans ce projet afin de cartographier annuellement les densités de populations de moustiques pour le sud du Québec représente pour l'INSPQ et l'ASPC un premier pas vers le développement d'un outil pratique d'aide à la décision. En effet, aucune méthodologie n'avait encore été développée dans la province pour être applicable à grande échelle et sur quatre espèces distinctes. Bien que nous ayons employé une méthodologie relativement simple et des variables environnementales génériques pour les appliquer aux différentes espèces, nos résultats sont encourageants et nos modèles offrent la possibilité de prévoir l'impact des changements climatiques et environnementaux sur la dispersion des populations de moustiques.

Afin de poursuivre cette recherche, un autre projet pourrait être entrepris afin de comparer la localisation des cas de VNO ou d'EEEV déclarés avec les *hot spots* de notre cartographie prédictive. Il faut distinguer ces *hot-spots* qui couvrent au total de grandes surfaces régionales, avec les cas de maladie avérés chez les hôtes ou avec les taux d'infection chez les animaux qui font office de réservoirs. Les réservoirs sont des espèces, ici animales, porteuses de virus pathogènes que le moustique contracte en les piquant et qu'il transmet ensuite à ses hôtes, soit les humains ou d'autres animaux en fonction des espèces de moustiques. Par exemple, les réservoirs reconnus pour l'EEEV et le VNO sont les oiseaux et le réservoir principal de LACV sont les mammifères sauvages (Clements, 2012) comme les chevreuils par exemple. Dans cette nouvelle étude, notre projet pourrait ainsi faire le pont entre les réservoirs et les hôtes en comparant la variation spatiale des abondances annuelles des espèces de moustiques cibles avec celles des oiseaux porteurs du VNO par exemple, et la localisation des populations le plus à risque dans les milieux urbains, en intégrant des paramètres socio-économiques par exemple (Kotchi *et al.*, 2016; Tachiiri *et al.*, 2006).

D'autre part, la base de données de capture a été préparée en incluant plusieurs autres espèces. Quatre d'entre elles n'ont pas donné de résultats satisfaisants à l'étape de validation de la normalité des résidus et ont été écartées, soit *Ochlerotatus canadensis*, *Aedes cinereus*, *Ochlerotatus provocans*, *Anopheles quadrimaculatus*. Trois autres espèces ont également été abandonnées au début des analyses, car leur abondance n'était pas suffisante pour notre méthodologie (*Ochlerotatus communis*, *Ochlerotatus triseriatus* et *Culiseta melanura*). Ces données pourraient faire l'objet d'un projet de recherche basée sur une méthodologie similaire,

mais qui emploierait un modèle qui soit adapté aux distributions particulières à ces espèces. Ce type de recherche, au même titre que celle présentée dans ce mémoire, fait avancer la connaissance des moustiques à risque pour la santé publique et d'anticiper l'évolution probable de leur répartition sur le territoire québécois en offrant des outils capables de prédire ces évolutions.

La modélisation prédictive rencontre deux objectifs différents que l'on retrouve tous deux dans la littérature : d'une part un objectif à long terme qui offre, comme notre projet, d'envisager des projections basées sur les impacts anticipés du réchauffement climatique. Dans ce cas, si le modèle est conçu pour intégrer des variables d'occupation du sol, alors il convient de modéliser l'évolution du paysage sur la période correspondante. À cet effet, Doelman *et al.*, (2018) proposent une méthode pour modéliser l'évolution du paysage à long terme sous l'influence des différents scénarios de réchauffement climatique. Par ailleurs, Verburgh *et al.*, (2009) conseillent de prendre en compte les évolutions des services et fonctions du sol qui permettent d'anticiper plus précisément les changements dans le temps au sein des COS. D'autre part la modélisation peut remplir l'objectif à court terme de prédire les variations d'abondance au cours d'une saison afin de guider directement les interventions sur le terrain plutôt que les instances décisionnelles. Dans un cas comme dans l'autre, la base de données existe et elle est suffisamment riche pour soutenir d'autres projets de recherche. Je souhaite donc que ce mémoire ouvre la voie à davantage de travail sur cette base de données ou que de futures études viennent compléter nos résultats.

## 8. Références

- Addinsoft (2019) XLSTAT 2019: Data Analysis and Statistical Solution for Microsoft Excel. Paris, France.
- Albers M.A., Bradley T.J. (2011) On the evolution of saline tolerance in the larvae of mosquitoes in the genus *Ochlerotatus*. *Physiological and Biochemical Zoology*, vol. 84, n° 3, p. 258–267.
- Anderson J.F., Main A.J., Armstrong P.M., Ferrandino F.J. (2015) Arboviruses in North Dakota, 2003–2006. *The American journal of tropical medicine and hygiene*, vol. 92, n° 2, p. 377–393.
- Andreadis T.G., Anderson J.F., Tirrell-Peck S.J. (1998) Multiple isolations of eastern equine encephalitis and highlands J viruses from mosquitoes (Diptera: Culicidae) during a 1996 epizootic in southeastern Connecticut. *Journal of medical entomology*, vol. 35, n° 3, p. 296–302.
- Andreadis T.G., Anderson J.F., Armstrong P.M., Main A.J. (2008) Isolations of Jamestown Canyon virus (Bunyaviridae: Orthobunyavirus) from field-collected mosquitoes (Diptera: Culicidae) in Connecticut, USA: a ten-year analysis, 1997–2006. *Vector-Borne and Zoonotic Diseases*, vol. 8, n° 2, p. 175–188.
- Andreadis T.G., Armstrong P.M., Anderson J.F., Main A.J. (2014) Spatial-temporal analysis of Cache Valley virus (Bunyaviridae: Orthobunyavirus) infection in anopheline and culicine mosquitoes (Diptera: Culicidae) in the northeastern United States, 1997–2012. *Vector-Borne and Zoonotic Diseases*, vol. 14, n° 10, p. 763–773.
- Aubry P., Gaüzère B.A. (2017) Arboviroses tropicales, actualités 2015. *in* Médecine tropicale. Diplôme de médecine tropicale des pays de l’Océan Indien, Saint-Denis, Île de la Réunion, <http://medecinetropicale.free.fr/cours/arboviroses.pdf>.
- Berry R.L., Parsons M.A., Lalonde-Weigert B.J., Lebio J. (1986) *Aedes Canadensis*. A Vector of La Crosse virus (California serogroup) in Ohio. *Journal of the American Mosquito Control Association-Mosquito News*, vol. 2, n° 1, p. 73–78.
- Bodner D., LaDeau S.L., Leisnham, P.T. (2018) Relationships Among Immature-Stage Metrics and Adult Abundances of Mosquito Populations in Baltimore, MD. *Journal of medical entomology*, vol. 56, n° 1, p. 192–198.
- Boscoe F.P. (Ed.). (2013) Geographic health data: fundamental techniques for analysis. CABI. Chap. 7, Spatial Cluster Analysis, Chen, J.
- Bosak P.J., Crans W.J. (2002) The structure and function of the larval siphon and spiracular apparatus of *Coquillettidia perturbans*. *Journal of the American Mosquito Control Association-Mosquito News*, vol. 18, n° 4, p. 280–283.
- Bouden M., Moulin B., Gosselin P. (2008) The geosimulation of West Nile virus propagation: a multi-agent and climate sensitive tool for risk management in public health. *International Journal of Health Geographics*, vol. 7, n° 1, p. 1.
- Brownstein J.S., Rosen H., Purdy D., Miller J. R., Merlino M., Mostashari F., Fish, D. (2002) Spatial analysis of West Nile virus: rapid risk assessment of an introduced vector-borne zoonosis. *Vector Borne and Zoonotic Diseases*, vol. 2, n° 3, p. 157–164.
- Brunkard J.M., López J.L.R., Ramirez J., Cifuentes E., Rothenberg S.J., Hunsperger E.A., Chester G.M., Brussolo R.M., Villarreal N.A., Haddad B.M. (2007) Dengue fever seroprevalence and risk factors, Texas—Mexico border, 2004. *Emerging infectious diseases*, vol. 13, n° 10, p. 1477.
- Canards Illimités (2018) Notre travail / Secteurs d’intérêt — Milieux humides. *in* Canards Illimités Canada, Pour l’eau, Pour la faune, Pour tous, Stonewall, Manitoba, Canada, <http://www.canards.ca/notre-travail/milieux-humides/>.
- Chen C.C., Epp T., Jenkins E., Waldner C., Curry P.S., Soos C. (2013) Modeling monthly variation of *Culex tarsalis* (Diptera: Culicidae) abundance and West Nile virus infection rate in the Canadian Prairies. *International journal of environmental research and public health*, vol. 10, n° 7, p. 3033–3051.

- Cleckner H.L., Allen T.R., Bellows A.S. (2011) Remote sensing and modeling of mosquito abundance and habitats in Coastal Virginia, USA. *Remote Sensing*, vol. 3, n° 12, p. 2663–2681.
- Clements A.N. (1999) *The biology of mosquitoes: sensory, reception and behavior*, Wallingford: CABI Publishing, vol. 2, 740 p.
- Clements A.N. (2000) *The biology of mosquitoes, Development, nutrition and reproduction*, Wallingford: CABI Publishing, vol. 1, 511 p.
- Clements A.N. (2012) *The biology of mosquitoes, Transmission of viruses and interactions with bacteria*, Wallingford: CABI Publishing vol. 3, 571 p.
- Clements A.C., Pfeiffer D.U. (2009) Emerging viral zoonoses: frameworks for spatial and spatiotemporal risk assessment and resource planning. *The veterinary journal*, vol. 182, n° 1, p. 21–30.
- Crans W.J. (2004) A classification system for mosquito life cycles: life cycle types for mosquitoes of the northeastern United States. *Journal of Vector Ecology*, vol. 29, p. 1–10.
- Diuk-Wasser M.A., Brown H.E., Andreadis T.G., Fish D. (2006) Modeling the spatial distribution of mosquito vectors for West Nile virus in Connecticut, USA. *Vector-Borne and Zoonotic Diseases*, vol. 6, n° 3, p. 283–295.
- Doelman J.C., Stehfest E., Tabeau A., Van Meijl H., Lassaletta L., Gernaat D.E.H.J., Hermans M., Daioglou V., Biemans H., van der Sluis S., van Vuuren D.P (2018) Exploring SSP land-use dynamics using the IMAGE model: Regional and gridded scenarios of land-use change and land-based climate change mitigation. *Global Environmental Change*, vol. 48, p. 119–135.
- Dohoo I.R., Martin W., Stryhn H. (2003) *Veterinary epidemiologic research*, 2<sup>nd</sup> Edition. Charlottetown, Canada: AVC Incorporated.
- Edillo F., Kiszewski A., Manjourides J., Pagano M., Hutchinson M., Kyle A., Arias J., Gaines D. Lampman R., Novak R., Foppa I., Lubelczyk C., Smith R., Moncayo A., Spielman A. and The *Culex pipiens* Working Group (2009) Effects of latitude and longitude on the population structure of *Culex pipiens* sl, vectors of West Nile virus in North America. *The American journal of tropical medicine and hygiene*, vol. 81, n° 5, p. 842–848.
- ESRI (2019) ArcGIS Desktop: Release 10.6. Environmental Systems Research Institute. Redlands, California, United-States.
- Gardner A.M., Lampman R.L., Muturi E.J. (2014) Land use patterns and the risk of West Nile virus transmission in central Illinois. *Vector-Borne and Zoonotic Diseases*, vol. 14, n° 5, p. 338–345.
- Goddard L.B., Roth A.E., Reisen W.K., Scott T.W. (2002) Vector competence of California mosquitoes for West Nile virus. *Emerging infectious diseases*, 2002, vol. 8, n° 12, p. 1385.
- Gouvernement du Canada (2018a) Surveillance du virus du Nil Occidental. *in* Gouvernement du Canada, Ottawa. <https://www.canada.ca/fr/sante-publique/services/maladies/virus-nil-occidental/surveillance-virus-nil-occidental.html>.
- Gouvernement du Canada (2018 b) Surveillance du virus Zika. *in* Gouvernement du Canada, Ottawa. <https://www.canada.ca/fr/sante-publique/services/maladies/virus-zika/surveillance-virus-zika.html>.
- Gouvernement du Canada (2018c) Données historiques. *in* Gouvernement du Canada, Ottawa. [http://climat.meteo.gc.ca/historical\\_data/search\\_historic\\_data\\_f.html](http://climat.meteo.gc.ca/historical_data/search_historic_data_f.html).
- Gouvernement du Canada (2018d) Pour les professionnels de la santé : virus du Nil occidental. *in* Gouvernement du Canada, Ottawa. <https://www.canada.ca/fr/sante-publique/services/maladies/virus-nil-occidental/pour-professionnels-sante-traitant-patients-infectes-virus-nil-occidental.html>.
- Gouvernement du Canada (2019) Couverture terrestre. *in* Gouvernement du Canada, Ottawa. <https://ouvert.canada.ca/data/fr/dataset/d1fc6010-e2e7-401a-8dc1-544cd2ac0b03>.
- Hamer G.L., Anderson T.K., Donovan D.J., Brawn J.D., Krebs B.L., Gardner A.M., Riuz M.O., Brown W.M., Kitron U.D., Newman C.M., Goldberg, T.L. (2014) Dispersal of adult *Culex* mosquitoes in an urban West Nile virus hotspot: a mark-capture study incorporating stable isotope enrichment of natural larval habitats. *PLoS neglected tropical diseases*, vol. 8, n° 3, p. 1–7.

- Han J., Zhu L., Kulldorff M., Hostovich S., Stinchcomb D.G., Tatalovich Z., Riedel-Lewis D., Feuer E.J. (2016) Using Gini coefficient to determining optimal cluster reporting sizes for spatial scan statistics. *International journal of health geographics*, vol. 15, n° 1, p. 27.
- Harris M.C., Yang F., Jackson D.M., Dotseth E.J., Paulson S L., Hawley D.M. (2015) La Crosse virus field detection and vector competence of *Culex* mosquitoes . *The American journal of tropical medicine and hygiene*, vol. 93, n° 3, p. 461-467.
- Hongoh V., Berrang-Ford L., Scott M.E., Lindsay, L.R. (2012) Expanding geographical distribution of the mosquito, *Culex pipiens*, in Canada under climate change. *Applied geography*, vol. 33, p. 53–62.
- Hoover K.C., Barker C.M. (2016) West Nile virus, climate change, and circumpolar vulnerability. *Wiley Interdisciplinary Reviews: Climate Change*, vol. 7, n° 2, p. 283–300.
- Île Saint-Bernard (2018) Héritage Saint-Bernard Entre ciel, terre et eau. *in* Bienvenu sur l'île Saint-Bernard, Châteauguay, <https://www.ilesaintbernard.com/propos>.
- Jacob B.G., Gu W., Caamano E.X., Novak R.J. (2009a) Developing operational algorithms using linear and non-linear squares estimation in Python® for the identification of *Culex pipiens* and *Culex restuans* in a mosquito abatement district (Cook County, Illinois, USA). *Geospatial health*, vol. 3, n° 2, p. 157–176.
- Jacob B.G., Lampman R.L., Ward M.P., Muturi E.J., Morris J.A., Caamano E.X., Novak R.J. (2009b) Geospatial variability in the egg raft distribution and abundance of *Culex pipiens* and *Culex restuans* in Urbana-Champaign, Illinois. *International Journal of Remote Sensing*, vol. 30, n° 8, p. 2005–2019.
- Johnson T.L., Haque U., Monaghan A.J., Eisen L., Hahn M.B., Hayden M.H., Savage H.M., McAllister J., Mutebi J.P., Eisen R. (2017) Modeling the environmental suitability for *Aedes* (*Stegomyia*) *aegypti* and *Aedes* (*Stegomyia*) *albopictus* (Diptera: Culicidae) in the Contiguous United States. *Journal of Medical Entomology*.
- Kalluri, S., Gilruth, P., Rogers, D., Szczur, M. (2007) Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: a review. *PLoS Pathog*, vol. 3, n° 10, p. e116.
- Karki S., Hamer G.L., Anderson T.K., Goldberg T.L., Kitron U.D., Krebs B.L., Walker D.W, Ruiz M.O. (2016) Effect of trapping methods, weather, and landscape on estimates of the *Culex* vector mosquito abundance. *Environmental health insights*, vol. 10, p. 93–103. EHI-S33384.
- Kotchi S.O., Brazeau S., Ludwig A., Aube G., Berthiaume P. (2016) Earth Observation and Indicators Pertaining to Determinants of Health-An Approach to Support Local Scale Characterization of Environmental Determinants of Vector-Borne Diseases. *Living Planet Symposium*, vol. 740, p. 9.
- Lebl K., Brugger K., Rubel F. (2013) Predicting *Culex pipiens/restuans* population dynamics by interval lagged weather data. *Parasites & vectors*, vol. 6, n° 1, p. 1.
- Liu H., Weng Q. (2012) Enhancing temporal resolution of satellite imagery for public health studies: A case study of West Nile Virus outbreak in Los Angeles in 2007. *Remote Sensing of Environment*, vol. 117, p. 57–71.
- Main A.J., Brown S.E., Wallis R.C., Elston, J. (1979) Arbovirus surveillance in Connecticut. II. California serogroup [*Aedes* species, insect vectors]. *Mosquito News*, vol. 39, n° 3, p. 552–559.
- Maire A., Aubin A. (1980) Les moustiques du Québec (Diptera : culicidae). *Essai de synthèse écologique. Mémoires de la société entomologique du Québec*, n° 6, p. 109.
- Marcantonio M., Rizzoli A., Metz M., Rosà R., Marini G., Chadwick E., Neteler M. (2015) Identifying the environmental conditions favouring West Nile Virus outbreaks in Europe. *PloS one*, vol. 10, n° 3, p. e0121158.
- Murdock C.C., Olival K.J., Perkins S.L. (2014) Molecular identification of host feeding patterns of snow-melt mosquitoes (Diptera: Culicidae): potential implications for the transmission ecology of Jamestown Canyon virus. *Journal of medical entomology*, vol. 47, n° 2, p. 226–229.
- Ministère de la Santé et des Services sociaux du Québec (2018) Surveillance des maladies d'intérêt transmises par des moustiques au Québec. *in* Gouvernement du Québec, Québec,



<http://www.msss.gouv.qc.ca/professionnels/zoonoses/surveillance-des-maladies-d-interet-transmises-par-des-moustiques-au-quebec/>.

- Organisation mondiale de la Santé (2017) Accueil. in Organisation mondiale de la santé, Genève, <http://www.who.int/mediacentre/factsheets/fs387/fr/>.
- Ozdenerol E., Bialkowska-Jelinska E., Taff G.N. (2008) Locating suitable habitats for West Nile Virus-infected mosquitoes through association of environmental characteristics with infected mosquito locations: a case study in Shelby County, Tennessee. *International Journal of Health Geographics*, vol. 7, n° 1, p. 1.
- Ozdenerol, E., Taff G.N., Akkus C. (2013) Exploring the spatio-temporal dynamics of reservoir hosts, vectors, and human hosts of West Nile virus: a review of the recent literature. *International journal of environmental research and public health*, 2013, vol. 10, n° 11, p. 5399–5432.
- PCI (2019) PCI Geomatics, Geomatica: Makham, Ontario, Canada.
- Platt R.B., Love G.J., Williams E.L. (1958) A positive correlation between relative humidity and the distribution and abundance of *Aedes vexans*. *Ecology*, vol. 39, n° 1, p. 167-169.
- Pradel J., Rey D., Foussadier R., Bicoud D. (2007) Étude écologique des moustiques (Diptera, culicidae) – vecteurs potentiels d’arboviroses – dans la région Rhône-Alpes. *Bulletin de l’Association pour l’Étude de l’Épidémiologie des Maladies Animales, Épidémiologie et santé animale*, vol. 51, p. 81-94.
- Pritchard E.J., Lewis D.J., Rau M.E. (2010) Seasonal abundance and distribution of mosquito species (Culicidae) in southwestern Québec and their potential role as vectors of West Nile Virus (Doctoral dissertation, McGill University), 174 p.
- Ressources naturelles Canada (2017) Sélection de produits Géogratis. in Gouvernement du Canada, Ottawa <http://geogratis.gc.ca/api/fr/nrcan-rncan/ess-sst/?entry-type=link>.
- Rocheleau J.P. (2016) Les caractéristiques environnementales du risque d’exposition aux arbovirus au Québec. Thèse de doctorat. Faculté de médecine vétérinaire de l’université de Montréal, Canada.
- Ripoche M., Campagna C., Ludwig A., Ogden N.H., Leighton P.A. (2019) Short-term forecasting of daily abundance of West Nile virus vectors *Culex pipiens-restuans* and *Aedes vexans* based on weather conditions and larvicide use in southern Québec (Canada). *Journal of Medical Entomology*.
- Ruiz M.O., Chaves L.F., Hamer G.L., Sun T., Brown W.M., Walker E.D., Haramis L., Goldberg T.L., Kitron, U. D. (2010) Local impact of temperature and precipitation on West Nile virus infection in *Culex* species mosquitoes in northeast Illinois, USA. *Parasites & vectors*, vol. 3, n° 1, p. 19.
- Sakamoto W. (2018) Bias-reduced marginal Akaike information criteria based on a Monte Carlo method for linear mixed-effects models. *Scandinavian Journal of Statistics*.
- Samy A.M., Thomas S.M., Wahed A.A.E., Cohoon K.P., Peterson A.T. (2016) Mapping the global geographic potential of Zika virus spread. *Memórias do Instituto Oswaldo Cruz*, vol. 111, n° 9, p. 559–560.
- Satscan (2018) Home. in SaTScan, Software for the spatial, temporal, and space-time scan statistics. <https://www.satscan.org/>.
- Sekla L.H., Stackiw W., Brust R.A. (1980) Arbovirus isolations from mosquitoes in Manitoba. *Mosquito News*, vol. 40, n° 3, p. 377–380.
- Shuai J, Buck P, Sockett P, Aramini, J, Pollari F. (2006) A GIS-driven integrated real-time surveillance pilot system for national West Nile virus dead bird surveillance in Canada. *International Journal of Health Geographics*, vol. 5, n° 1, p. 1.
- Srihongse S., Woodall J.P., Grayson M.A., Deibel R., Bast T.F., Morris C.D., Bosler E.M., Benach J.L., Howard J.J., Berlin J. (1980) Arboviruses in New York State: surveillance in arthropods and nonhuman vertebrates, 1972–1977. *Mosquito News*, vol. 40, n° 2, p. 269–276.
- Strickman D. (1980). Stimuli affecting selection of oviposition sites by *Aedes vexans* (Diptera: Culicidae): moisture. *Mosq News*, n° 40, p. 236–245.
- Strickman D. (1982) Stimuli affecting selection of oviposition sites by *Aedes vexans* (Diptera: Culicidae): Light. *Journal of Medical Entomology*, Vol. 19, n° 2, p. 181–184.

- Tachiiri K., Klinkenberg B., Mak S., Kazmi J. (2006) Predicting outbreaks: a spatial risk assessment of West Nile virus in British Columbia. *International Journal of Health Geographics*, vol. 5, n° 1, p. 1.
- Trawinski P.R., Mackay D.S. (2009) Spatial autocorrelation of West Nile virus vector mosquito abundance in a seasonally wet suburban environment. *Journal of geographical systems*, vol. 11, n° 1, p. 67–87.
- Trawinski P.R., Mackay D.S. (2010) Identification of environmental covariates of West Nile virus vector mosquito population abundance. *Vector-Borne and Zoonotic Diseases*, vol. 10, n° 5, p. 515–526.
- Turell M.J., Sardelis M.R., Dohm D.J., O'guinn M.L. (2001) Potential North American vectors of west Nile virus. *Annals of the New York Academy of Sciences*, vol. 951, n° 1, p. 317–324.
- Turell M.J., Dohm D.J., Sardelis M.R., O'guinn M.L., Andreadis T.G., Blow J.A. (2005) An update on the potential of North American mosquitoes (Diptera: Culicidae) to transmit West Nile virus. *Journal of medical entomology*, vol. 42, n° 1, p. 57–62.
- Verburg P.H., Van De Steeg J., Veldkamp A., Willemen L. (2009) From land cover change to land function dynamics: a major challenge to improve land characterization. *Journal of environmental management*, vol. 90, n° 3, p. 1327-1335.
- Verdonschot P.F.M., Besse-Lototskay A.A. (2014) Flight distance of mosquitoes (Culicidae): a metadata analysis to support the management of barrier zones around rewetted and newly constructed wetlands. *Limnologia-Ecology and Management of Inland Waters*, vol. 45, p. 69–79.
- Wagner R.J., DeJong C., Leung M.K., McLintock J., Iversen J.O. (1975) Isolations of California encephalitis virus from tundra mosquitoes. *Canadian journal of microbiology*, vol. 21, n° 4, p. 574–576.
- Wajnberg E. (2011) Introduction au Modèle Linéaire Généralisé (Generalized linear Model; GML). Cours UE7. Université de Nice-Sophia-Antipolis.
- Walker E.D., Grayson M.A., Edman J.D. (1993) Isolation of Jamestown Canyon and Snowshoe Hare Viruses (California Serogroup) from Aedes Mosquitoes in Western Massachusetts. *Journal-american mosquito control association*, vol. 9, p. 131–131.
- Wimberly M.C., Lamsal A., Giacomo P., Chuang T.W. (2014) Regional variation of climatic influences on West Nile virus outbreaks in the United States. *The American journal of tropical medicine and hygiene*, vol. 91, n° 4, p. 677–684.
- Wang G., Minnis R.B., Belant J.L., Wax C.L. (2010) Dry weather induces outbreaks of human West Nile virus infections. *BMC infectious diseases*, vol. 10, n° 1, p. 1.
- Wang J., Ogden N.H., Zhu H. (2011) The impact of weather conditions on Culex pipiens and Culex restuans (Diptera: Culicidae) abundance: a case study in Peel region. *Journal of medical entomology*, vol. 48, n° 2, p. 468–475.
- Wang X., Wang J., Russell C., Proctor P., Bello R., Higuchi K., Zhu H. (2014) Clustering of the abundance of West Nile virus vector mosquitoes in Peel Region, Ontario, Canada. *Environmental and ecological statistics*, vol. 21, n° 4, p. 651–666.
- Ward R.D., Darsie R.F. (2005) Identification and Geographical Distribution of the Mosquitoes of North America, North of Mexico, 416 p. University Press of Florida, USA.
- Wood D.M., Dang P.T., Ellis R.A. (1979) The insects and arachnids of Canada. Part 6. The mosquitoes of Canada. Diptera: Culicidae. Canadian Government Publishing Centre.
- Yoo E.H. (2014) Site-specific prediction of West Nile virus mosquito abundance in Greater Toronto Area using generalized linear mixed models. *International Journal of Geographical Information Science*, vol. 28, n° 2, p. 296–313.
- Yoo E.H., Chen D., Diao C., Russell C. (2016) The effects of weather and environmental factors on West Nile virus mosquito abundance in greater Toronto area. *Earth Interactions*, vol. 20, n° 3, p. 1–22.
- Zuur A.F., Ieno E.N., Elphick C.S. (2010) A protocol for data exploration to avoid common statistical problems. *Methods in ecology and evolution*, vol. 1, n° 1, p. 3–14.

---

ANNEXE

---

### Note sur l'autocorrélation spatiale

Comme pour l'étude de dispersion, l'autocorrélation spatiale mérite une analyse. En effet, les pièges sont majoritairement localisés en milieu urbain ou périurbain, ce qui induit *de facto* un biais dans la modélisation. Le I de Moran illustré en fonction de la distance (figure 1) a mis en évidence l'autocorrélation spatiale présente étant donnée la distribution des pièges sur le territoire. Cependant, il a été décidé de ne pas l'inclure dans les prochaines étapes pour des raisons de faisabilité dans les temps. Ce sera donc une amélioration possible à apporter aux modèles dans le futur.

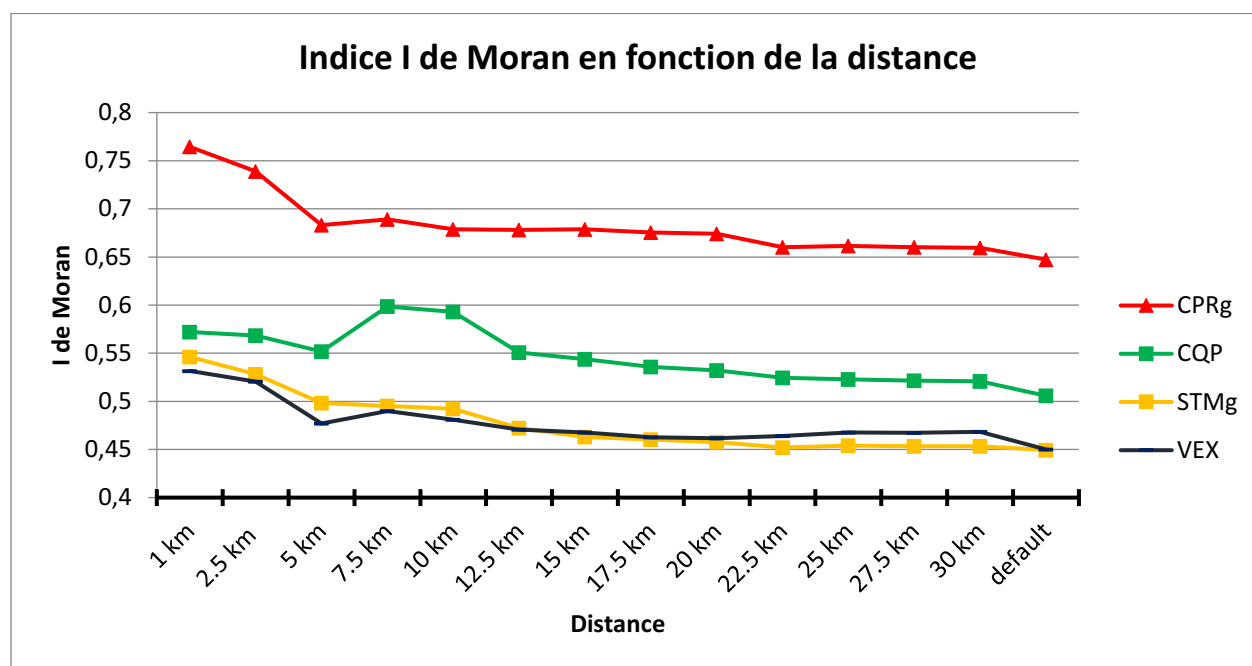


Figure A1 : Indice I de Moran en fonction de la distance pour les pièges des quatre espèces à l'étude.

### Note sur l'étude de dispersion des espèces au fil du temps

Dans le cadre d'une modélisation prédictive liée aux conditions favorables au développement des moustiques, face au contexte de changement climatique, on peut s'attendre à ce que leurs zones de distribution évoluent sur le territoire avec le temps (Samy *et al.*, 2016). Ainsi, pour que cette évolution puisse être prise en compte dans la modélisation, il faut l'évaluer et la quantifier. Le logiciel SaTScan (Satscan, 2018) est particulièrement adapté à ce type d'étude de dispersion spatio-temporelle. La confrontation de deux méthodologies différentes, décrites dans deux articles scientifiques distincts (Boscoe, 2013 ; Han *et al.*, 2016), a permis de valider que la méthode employée est efficace. La méthode a été appliquée à quatre espèces de moustiques au cours des treize dernières années. Cependant, aucune des quatre espèces testées ne semble avoir étendu son territoire d'une façon particulièrement significative, ni en direction ni en superficie. La période d'étude n'était peut-être pas suffisante pour démontrer une influence des changements climatiques, mais en l'absence de résultat tangible, la dispersion temporelle des espèces de moustiques n'est pas prise en compte dans cette modélisation.

## Phase exploratoire – Étape 1 – Cartographie de la répartition des pièges (paragraphe 5.1.1)

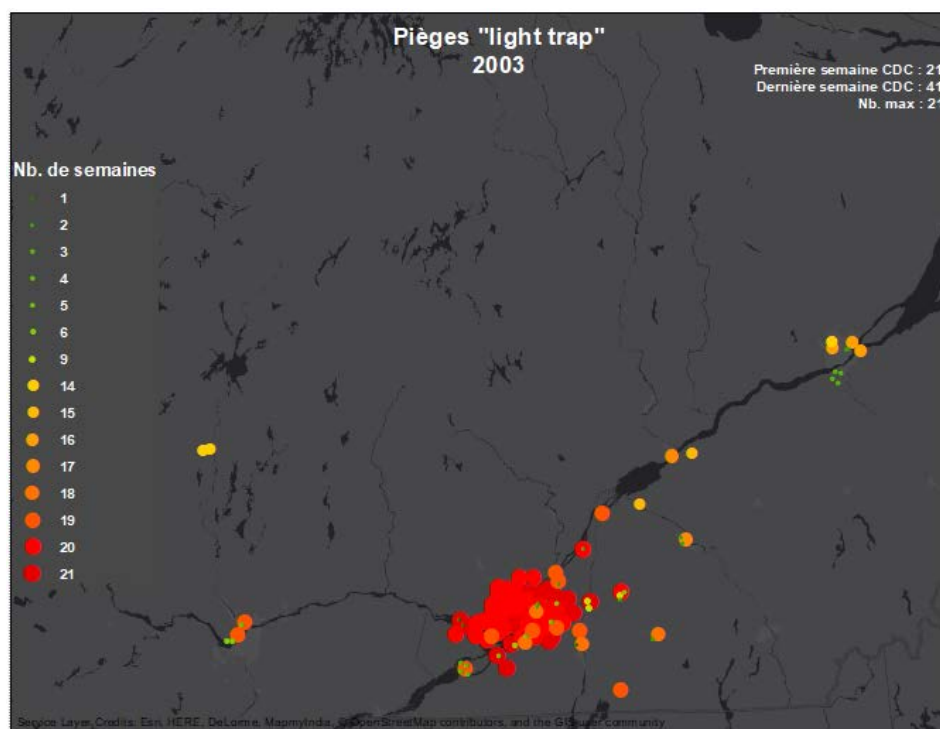


Figure A1 : Répartition des pièges selon leur durée de fonctionnement pour l'année d'entraînement 2003.

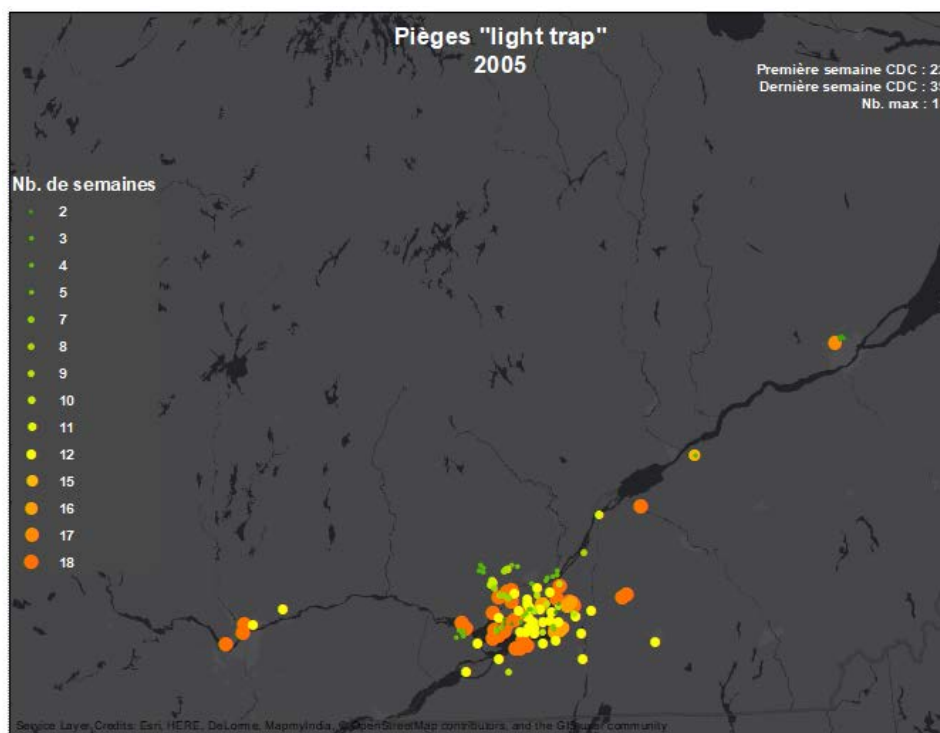
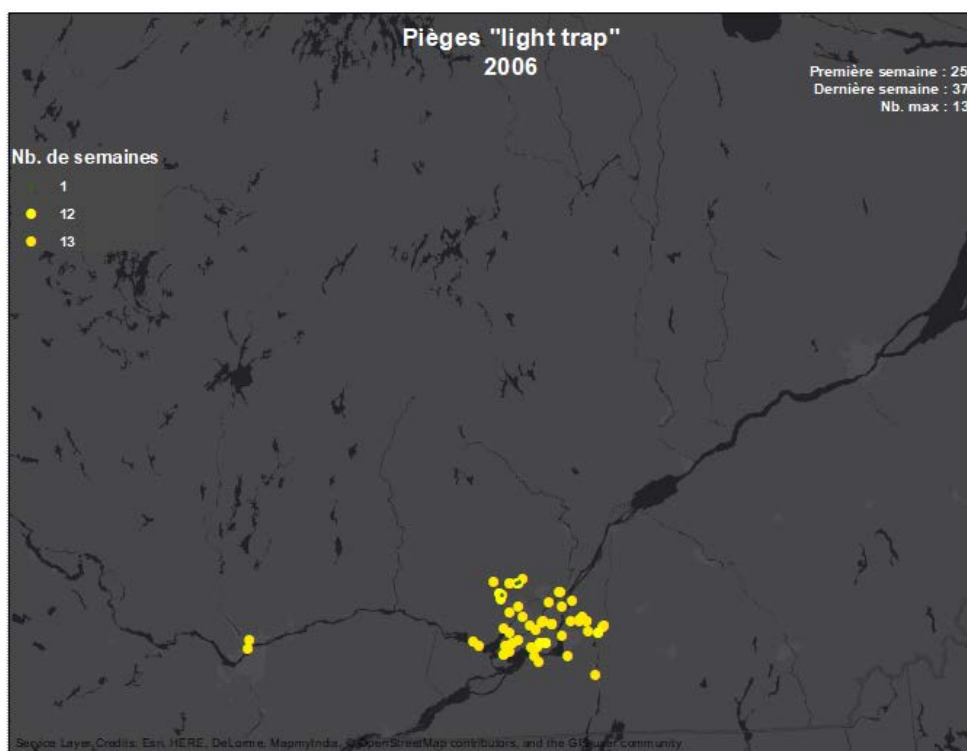
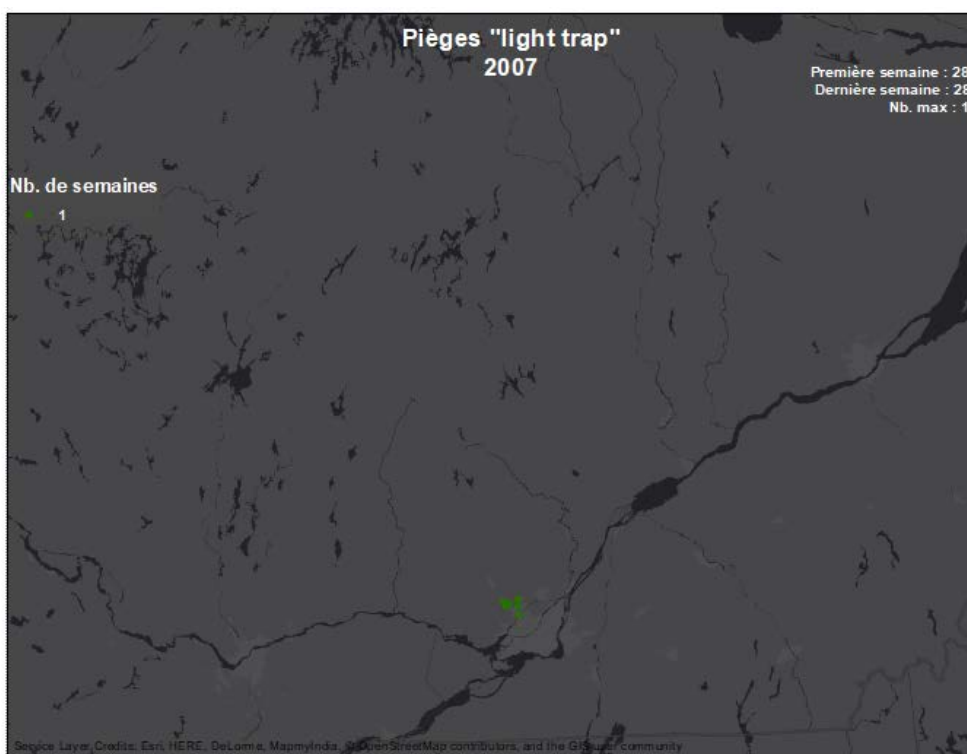


Figure A2 : Répartition des pièges selon leur durée de fonctionnement pour l'année d'entraînement 2005.



*Figure A3 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2006.*



*Figure A4 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2007.*

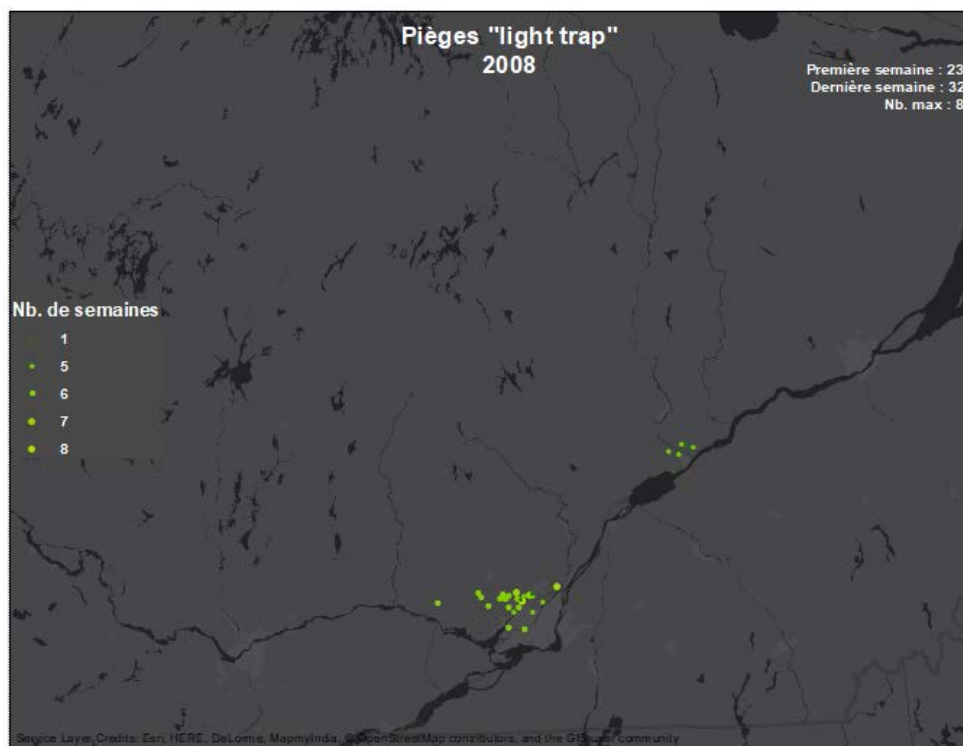


Figure A5 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2008.

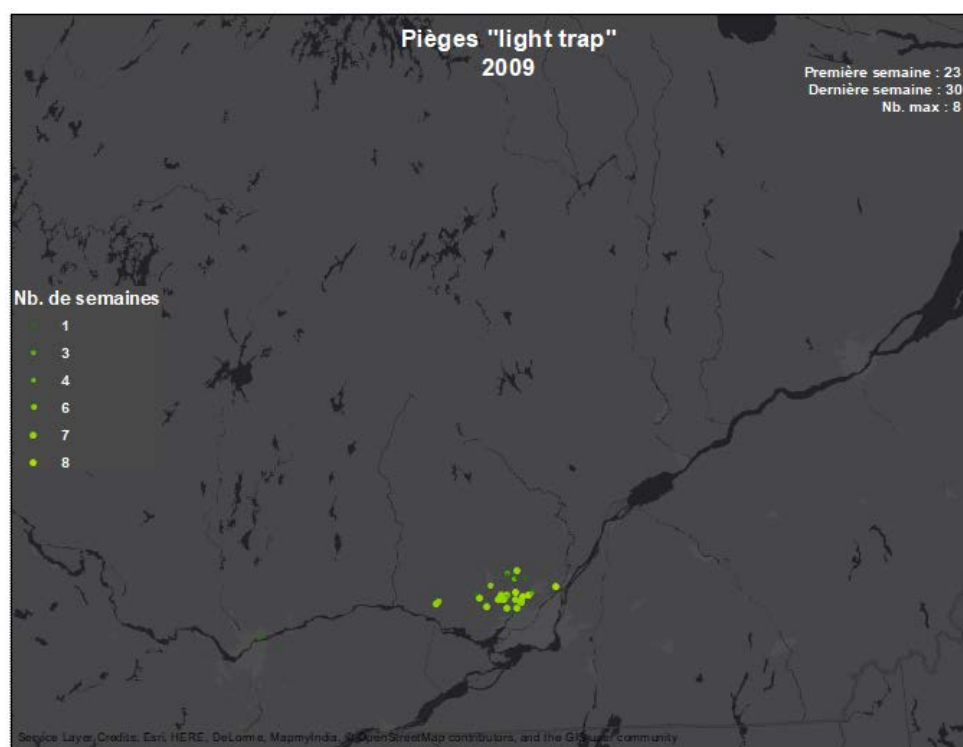


Figure A6 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2009.



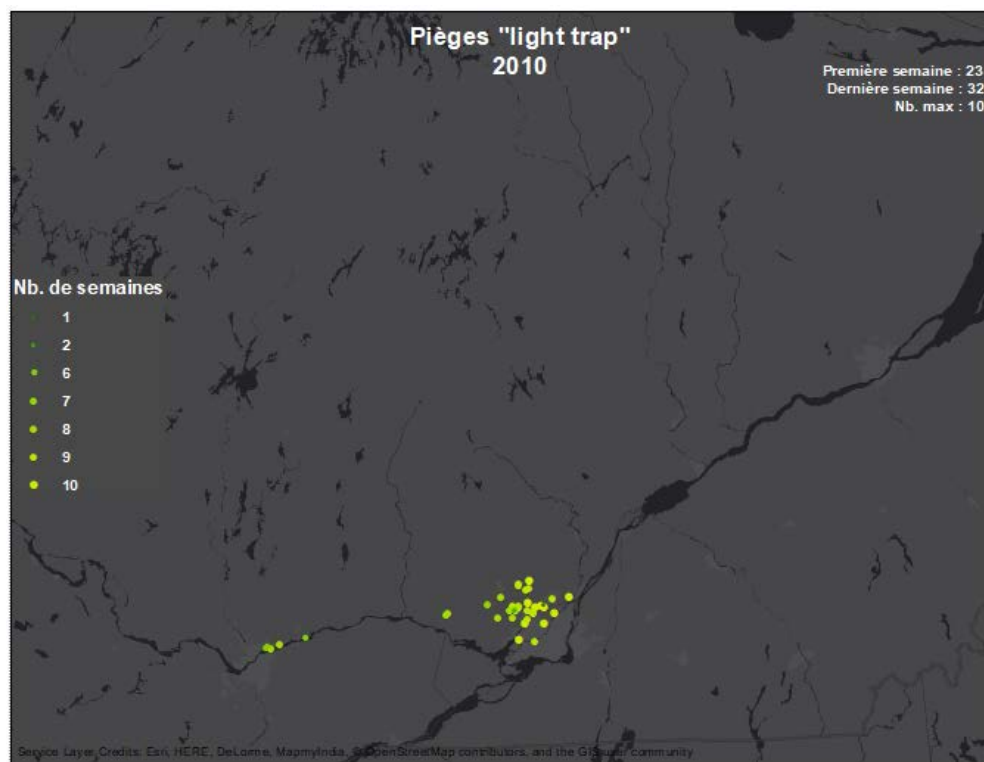


Figure A7 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2010.

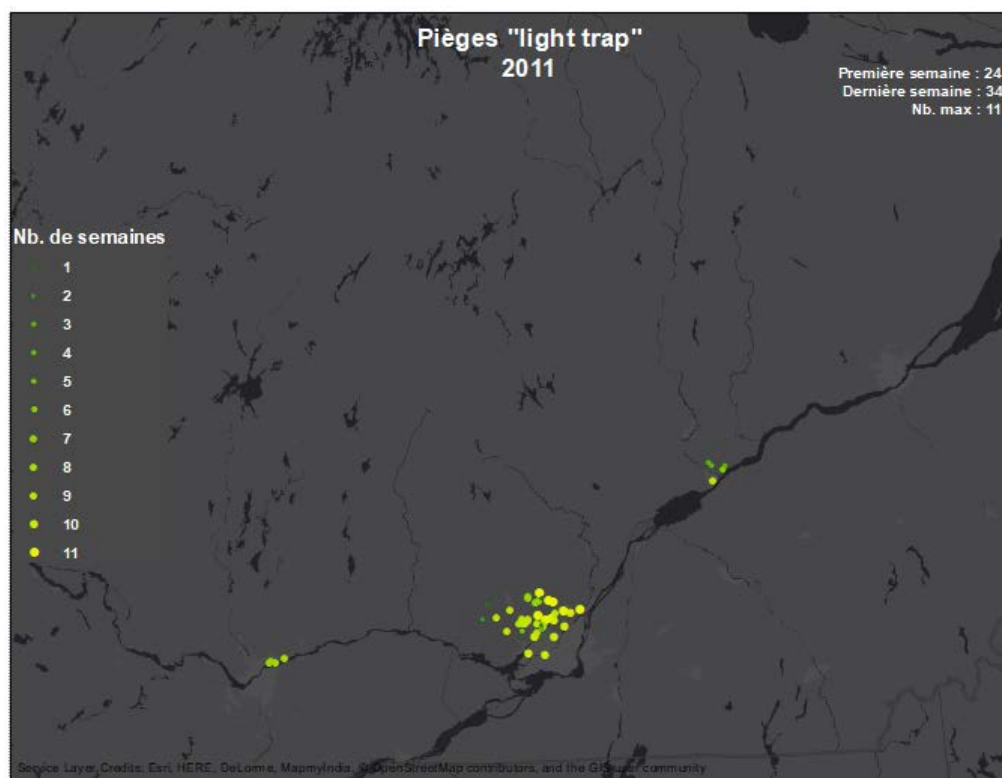


Figure A8 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2011.



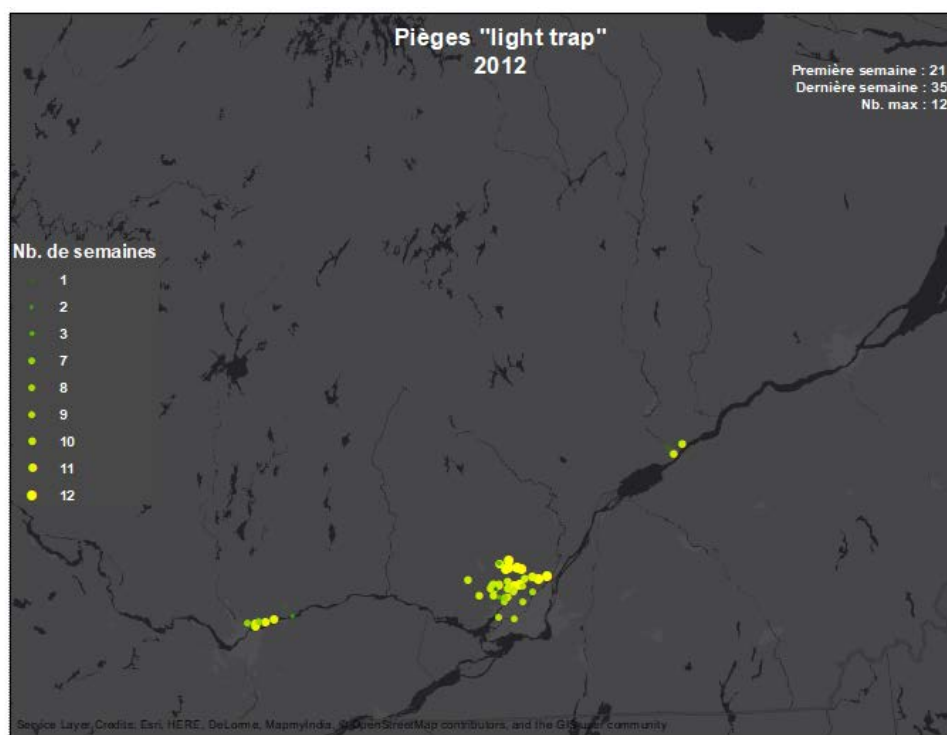


Figure A9 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2012.

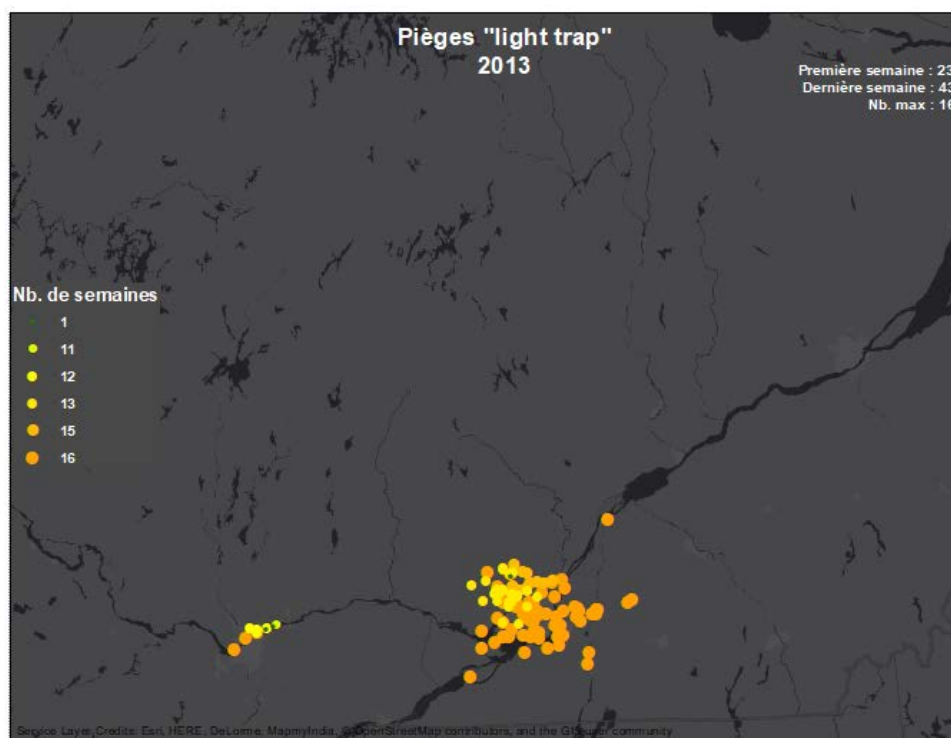
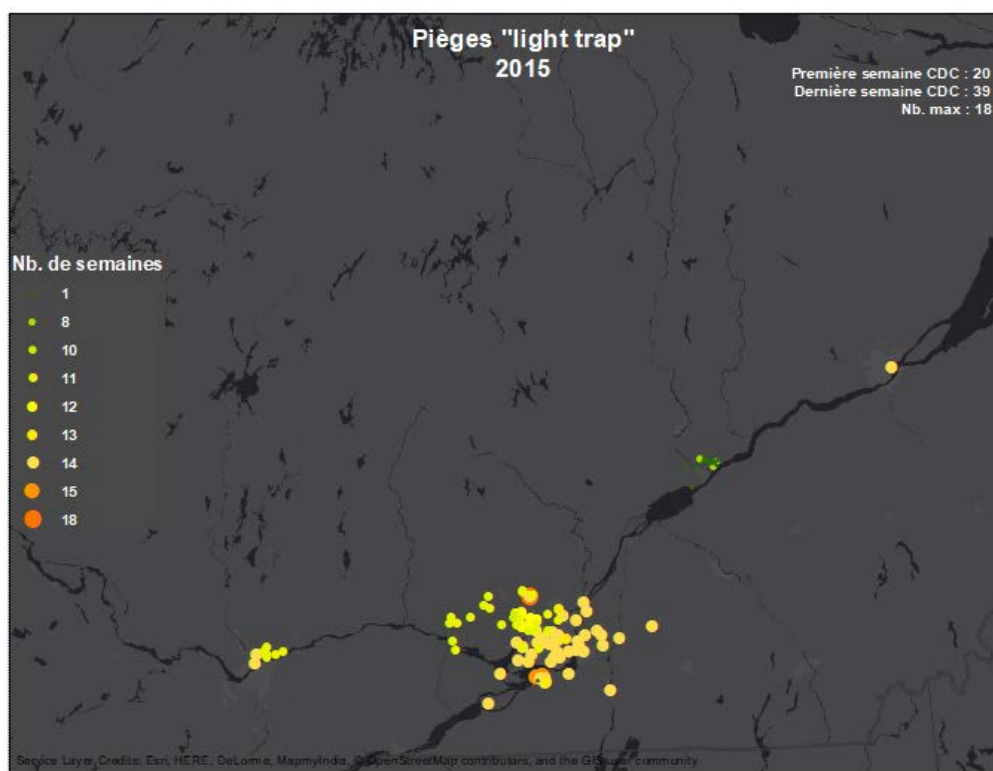
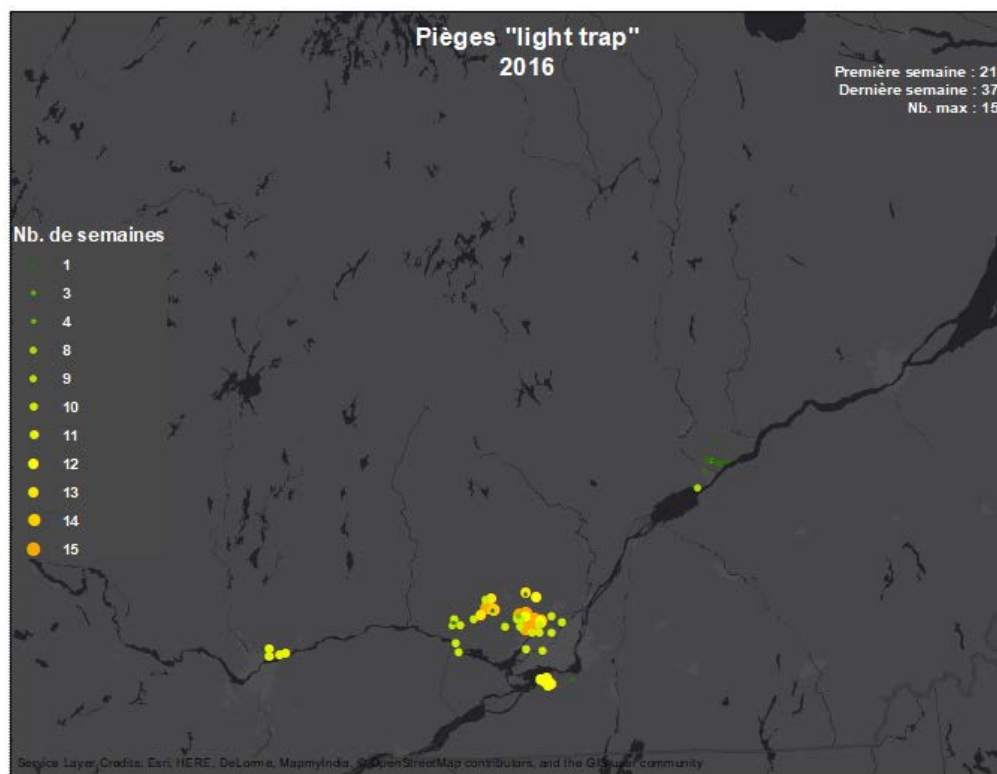


Figure A10 : Répartition des pièges selon leur durée de fonctionnement pour l'année d'entraînement 2013.



*Figure A11 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2015.*



*Figure A12 : Répartition des pièges selon leur durée de fonctionnement pour l'année rejetée 2016.*

## Phase de modélisation – Étape 2 – Régressions univariées (paragraphe 5.2.2)

Tableau A1 : Résultats des régressions univariées des variables environnementales pour chaque espèce et abréviations correspondantes.

Source (et abréviation)		Espèce	Valeur	Erreur standard	DDL	t	Pr >  t	Borne inférieure (95 %)	Borne supérieure (95 %)
Eau	Eau	CQP	-0,002	0,009	61	-0,25	<b>0,802</b>	-0,019	0,015
		VEX	-0,011	0,004	61	-2,91	<b>0,005</b>	-0,018	-0,003
		CPRg	-0,022	0,012	61	-1,90	<b>0,062</b>	-0,046	0,001
		STMg	0,059	0,034	61	1,74	<b>0,087</b>	-0,008	0,125
Sol nu	SolNu	CQP	0,045	0,119	61	0,38	<b>0,704</b>	-0,190	0,280
		VEX	-0,003	0,053	61	-0,05	<b>0,961</b>	-0,108	0,103
		CPRg	-0,053	0,165	61	-0,32	<b>0,748</b>	-0,379	0,272
		STMg	0,365	0,470	61	0,78	<b>0,441</b>	-0,563	1 293
Urbain	Urbain	CQP	-0,022	0,004	61	-5,35	<b>&lt; 0,0001</b>	-0,030	-0,014
		VEX	-0,003	0,002	61	-1,42	<b>0,161</b>	-0,007	0,001
		CPRg	0,012	0,006	61	1,86	<b>0,067</b>	-0,001	0,024
		STMg	-0,107	0,016	61	-6,76	<b>&lt; 0,0001</b>	-0,138	-0,076
Ensemble de milieux humides	MHtot	CQP	0,157	0,094	61	1,68	<b>0,098</b>	-0,028	0,342
		VEX	-0,017	0,043	61	-0,40	<b>0,690</b>	-0,102	0,068
		CPRg	-0,073	0,130	61	-0,56	<b>0,576</b>	-0,331	0,184
		STMg	0,624	0,370	61	1,68	<b>0,097</b>	-0,107	1,354
Agriculture	Agri	CQP	0,019	0,008	61	2,35	<b>0,022</b>	0,003	0,036
		VEX	0,016	0,003	61	4,57	<b>&lt; 0,0001</b>	0,009	0,023
		CPRg	0,057	0,010	61	5,50	<b>&lt; 0,0001</b>	0,037	0,078
		STMg	0,033	0,033	61	0,99	<b>0,325</b>	-0,033	0,098
Prairie	Prairie	CQP	0,021	0,019	61	1,09	<b>0,281</b>	-0,017	0,060
		VEX	0,000	0,009	61	0,00	<b>0,999</b>	-0,017	0,017
		CPRg	-0,054	0,027	61	-2,01	<b>0,048</b>	-0,106	-0,001
		STMg	0,169	0,076	61	2,21	<b>0,031</b>	0,018	0,319
Végétation	Vege	CQP	0,023	0,012	61	1,86	<b>0,068</b>	-0,001	0,047
		VEX	0,000	0,006	61	-0,06	<b>0,956</b>	-0,011	0,011
		CPRg	-0,024	0,017	61	-1,42	<b>0,162</b>	-0,058	0,010
		STMg	0,013	0,049	61	0,26	<b>0,795</b>	-0,084	0,110
Boisé	Boise	CQP	0,027	0,006	61	4,18	<b>&lt; 0,0001</b>	0,014	0,040
		VEX	0,003	0,003	61	1,03	<b>0,308</b>	-0,003	0,009
		CPRg	-0,035	0,009	61	-3,88	<b>0,000</b>	-0,054	-0,017
		STMg	0,150	0,024	61	6,23	<b>&lt; 0,0001</b>	0,102	0,197
Températures minimales annuelles	TNan	CQP	0,101	0,105	61	0,96	<b>0,338</b>	-0,106	0,308
		VEX	0,096	0,050	61	1,90	<b>0,062</b>	-0,004	0,195
		CPRg	0,185	0,142	61	1,30	<b>0,197</b>	-0,095	0,465
		STMg	-2,076	0,384	61	-5,41	<b>&lt; 0,0001</b>	-2,834	-1,319
Températures maximales annuelles	TXan	CQP	0,583	0,117	61	4,97	<b>&lt; 0,0001</b>	0,351	0,814
		VEX	0,421	0,053	61	7,97	<b>&lt; 0,0001</b>	0,317	0,525
		CPRg	0,793	0,163	61	4,86	<b>&lt; 0,0001</b>	0,471	1,116
		STMg	-1,219	0,488	61	-2,50	<b>0,015</b>	-2,182	-0,257

Températures moyennes annuelles	<b>TYan</b>	<b>CQP</b>	0,336	0,118	61	2,85	<b>0,006</b>	0,103	0,569
		<b>VEX</b>	0,258	0,056	61	4,64	<b>&lt; 0,0001</b>	0,148	0,368
		<b>CPRg</b>	0,469	0,161	61	2,91	<b>0,005</b>	0,151	0,787
		<b>STMg</b>	-1,930	0,454	61	-4,25	<b>&lt; 0,0001</b>	-2,825	-1,035
Précipitations totales annuelles	<b>Pan</b>	<b>CQP</b>	0,000	0,000	61	0,05	<b>0,963</b>	-0,001	0,001
		<b>VEX</b>	0,000	0,000	61	0,05	<b>0,963</b>	-0,001	0,001
		<b>CPRg</b>	0,000	0,001	61	0,05	<b>0,957</b>	-0,001	0,002
		<b>STMg</b>	-0,001	0,002	61	-0,55	<b>0,587</b>	-0,006	0,003
Températures minimales du mois de janvier	<b>TNjan</b>	<b>CQP</b>	-0,060	0,045	61	-1,35	<b>0,183</b>	-0,148	0,028
		<b>VEX</b>	-0,022	0,023	61	-0,97	<b>0,334</b>	-0,068	0,023
		<b>CPRg</b>	0,080	0,060	61	1,34	<b>0,186</b>	-0,038	0,199
		<b>STMg</b>	-0,964	0,153	61	-6,32	<b>&lt; 0,0001</b>	-1,265	-0,663
Températures minimales du mois de février	<b>TNfev</b>	<b>CQP</b>	0,096	0,033	61	2,90	<b>0,005</b>	0,031	0,162
		<b>VEX</b>	0,060	0,017	61	3,61	<b>0,001</b>	0,027	0,094
		<b>CPRg</b>	0,074	0,046	61	1,61	<b>0,113</b>	-0,017	0,165
		<b>STMg</b>	-0,686	0,126	61	-5,44	<b>&lt; 0,0001</b>	-0,934	-0,437
Températures minimales du mois de mars	<b>TNmar</b>	<b>CQP</b>	0,041	0,034	61	1,20	<b>0,236</b>	-0,027	0,109
		<b>VEX</b>	-0,016	0,018	61	-0,90	<b>0,372</b>	-0,052	0,019
		<b>CPRg</b>	0,006	0,046	61	0,13	<b>0,900</b>	-0,086	0,097
		<b>STMg</b>	-0,468	0,135	61	-3,45	<b>0,001</b>	-0,735	-0,200
Températures minimales du mois d'avril	<b>TNavr</b>	<b>CQP</b>	0,274	0,049	61	5,59	<b>&lt; 0,0001</b>	0,178	0,371
		<b>VEX</b>	0,174	0,024	61	7,33	<b>&lt; 0,0001</b>	0,127	0,221
		<b>CPRg</b>	0,209	0,074	61	2,84	<b>0,006</b>	0,064	0,354
		<b>STMg</b>	-0,736	0,213	61	-3,45	<b>0,001</b>	-1,157	-0,315
Températures minimales du mois de mai	<b>TNmai</b>	<b>CQP</b>	-0,355	0,080	61	-4,42	<b>&lt; 0,0001</b>	-0,513	-0,197
		<b>VEX</b>	-0,297	0,035	61	-8,60	<b>&lt; 0,0001</b>	-0,365	-0,229
		<b>CPRg</b>	-0,212	0,114	61	-1,86	<b>0,067</b>	-0,436	0,012
		<b>STMg</b>	-1,195	0,320	61	-3,74	<b>0,000</b>	-1,826	-0,564
Températures minimales du mois de juin	<b>TNjun</b>	<b>CQP</b>	0,119	0,047	61	2,51	<b>0,015</b>	0,026	0,213
		<b>VEX</b>	0,129	0,022	61	5,80	<b>&lt; 0,0001</b>	0,085	0,172
		<b>CPRg</b>	0,130	0,064	61	2,02	<b>0,048</b>	0,003	0,257
		<b>STMg</b>	-0,510	0,188	61	-2,71	<b>0,009</b>	-0,881	-0,139
Températures minimales du mois de juillet	<b>TNjul</b>	<b>CQP</b>	-0,019	0,137	61	-0,14	<b>0,889</b>	-0,289	0,251
		<b>VEX</b>	-0,045	0,063	61	-0,71	<b>0,480</b>	-0,170	0,080
		<b>CPRg</b>	-0,096	0,187	61	-0,51	<b>0,610</b>	-0,465	0,273
		<b>STMg</b>	-1,950	0,524	61	-3,72	<b>0,000</b>	-2,983	-0,917
Températures minimales du mois d'août	<b>TNaou</b>	<b>CQP</b>	-0,190	0,073	61	-2,62	<b>0,011</b>	-0,333	-0,047
		<b>VEX</b>	-0,079	0,038	61	-2,09	<b>0,041</b>	-0,153	-0,004
		<b>CPRg</b>	-0,077	0,101	61	-0,76	<b>0,449</b>	-0,275	0,122
		<b>STMg</b>	0,076	0,298	61	0,26	<b>0,799</b>	-0,512	0,664
Températures minimales du mois de septembre	<b>TNsep</b>	<b>CQP</b>	-0,044	0,079	61	-0,56	<b>0,578</b>	-0,201	0,112
		<b>VEX</b>	0,038	0,040	61	0,97	<b>0,335</b>	-0,040	0,117
		<b>CPRg</b>	-0,022	0,106	61	-0,21	<b>0,837</b>	-0,231	0,187
		<b>STMg</b>	0,254	0,312	61	0,82	<b>0,417</b>	-0,360	0,869
Températures minimales du mois de septembre de l'année précédente	<b>TNsep_1</b>	<b>CQP</b>	-0,246	0,078	61	-3,15	<b>0,003</b>	-0,400	-0,092
		<b>VEX</b>	-0,200	0,036	61	-5,61	<b>&lt; 0,0001</b>	-0,270	-0,129
		<b>CPRg</b>	-0,248	0,107	61	-2,32	<b>0,024</b>	-0,459	-0,037
		<b>STMg</b>	0,778	0,310	61	2,51	<b>0,015</b>	0,167	1,390

Températures minimales du mois d'octobre de l'année précédente	<b>TNoct_1</b> calculé dans R	<b>CQP</b>	0,046	0,056	61	0,82	<b>0,413</b>	-0,064	0,156
		<b>VEX</b>	-0,002	0,029	61	-0,08	<b>0,933</b>	-0,059	0,054
		<b>CPRg</b>	0,023	0,074	61	0,31	<b>0,754</b>	-0,123	0,170
		<b>STMg</b>	-1,263	0,196	61	-6,43	<b>&lt; 0,0001</b>	-1,651	-0,876
Températures minimales du mois de novembre de l'année précédente	<b>TNnov_1</b>	<b>CQP</b>	0,377	0,070	61	5,38	<b>&lt; 0,0001</b>	0,239	0,515
		<b>VEX</b>	0,203	0,035	61	5,74	<b>&lt; 0,0001</b>	0,133	0,273
		<b>CPRg</b>	0,189	0,107	61	1,77	<b>0,081</b>	-0,021	0,400
		<b>STMg</b>	0,315	0,308	61	1,03	<b>0,309</b>	-0,291	0,922
Températures minimales du mois de décembre de l'année précédente	<b>TNdec_1</b>	<b>CQP</b>	-0,203	0,033	61	-6,06	<b>&lt; 0,0001</b>	-0,269	-0,137
		<b>VEX</b>	-0,157	0,014	61	-11,33	<b>&lt; 0,0001</b>	-0,184	-0,129
		<b>CPRg</b>	-0,099	0,050	61	-1,97	<b>0,054</b>	-0,198	0,000
		<b>STMg</b>	0,024	0,149	61	0,16	<b>0,871</b>	-0,270	0,319
Températures maximales du mois de janvier	<b>TXjan</b>	<b>CQP</b>	-0,038	0,033	61	-1,18	<b>0,241</b>	-0,103	0,026
		<b>VEX</b>	-0,006	0,017	61	-0,34	<b>0,732</b>	-0,040	0,028
		<b>CPRg</b>	0,058	0,043	61	1,32	<b>0,190</b>	-0,028	0,143
		<b>STMg</b>	-0,636	0,112	61	-5,67	<b>&lt; 0,0001</b>	-0,857	-0,415
Températures maximales du mois de février	<b>TXfev</b>	<b>CQP</b>	0,206	0,030	61	6,98	<b>&lt; 0,0001</b>	0,148	0,265
		<b>VEX</b>	0,125	0,015	61	8,53	<b>&lt; 0,0001</b>	0,096	0,154
		<b>CPRg</b>	0,129	0,047	61	2,73	<b>0,008</b>	0,036	0,222
		<b>STMg</b>	-0,353	0,139	61	-2,54	<b>0,013</b>	-0,626	-0,079
Températures maximales du mois de mars	<b>TXmar</b>	<b>CQP</b>	0,125	0,059	61	2,12	<b>0,038</b>	0,009	0,242
		<b>VEX</b>	-0,004	0,032	61	-0,13	<b>0,894</b>	-0,067	0,058
		<b>CPRg</b>	0,042	0,081	61	0,52	<b>0,605</b>	-0,118	0,202
		<b>STMg</b>	-0,211	0,242	61	-0,87	<b>0,386</b>	-0,687	0,266
Températures maximales du mois d'avril	<b>TXavr</b>	<b>CQP</b>	0,246	0,035	61	6,93	<b>&lt; 0,0001</b>	0,176	0,316
		<b>VEX</b>	0,158	0,017	61	9,53	<b>&lt; 0,0001</b>	0,126	0,191
		<b>CPRg</b>	0,202	0,055	61	3,66	<b>0,001</b>	0,093	0,311
		<b>STMg</b>	-0,412	0,162	61	-2,55	<b>0,013</b>	-0,731	-0,093
Températures maximales du mois de mai	<b>TXmai</b>	<b>CQP</b>	-0,164	0,055	61	-2,99	<b>0,004</b>	-0,273	-0,056
		<b>VEX</b>	-0,168	0,025	61	-6,75	<b>&lt; 0,0001</b>	-0,217	-0,119
		<b>CPRg</b>	-0,046	0,076	61	-0,61	<b>0,544</b>	-0,196	0,104
		<b>STMg</b>	-0,253	0,221	61	-1,14	<b>0,258</b>	-0,690	0,184
Températures maximales du mois de juin	<b>TXjun</b>	<b>CQP</b>	0,189	0,048	61	3,95	<b>0,000</b>	0,095	0,284
		<b>VEX</b>	0,198	0,021	61	9,58	<b>&lt; 0,0001</b>	0,157	0,239
		<b>CPRg</b>	0,197	0,067	61	2,95	<b>0,004</b>	0,066	0,329
		<b>STMg</b>	0,181	0,197	61	0,92	<b>0,361</b>	-0,207	0,569
Températures maximales du mois de juillet	<b>TXjul</b>	<b>CQP</b>	0,390	0,107	61	3,64	<b>0,001</b>	0,179	0,602
		<b>VEX</b>	0,315	0,050	61	6,29	<b>&lt; 0,0001</b>	0,216	0,414
		<b>CPRg</b>	0,561	0,146	61	3,85	<b>0,000</b>	0,273	0,848
		<b>STMg</b>	-1,207	0,427	61	-2,83	<b>0,006</b>	-2,050	-0,365
Températures maximales du mois d'août	<b>TXaou</b>	<b>CQP</b>	0,104	0,066	61	1,58	<b>0,119</b>	-0,026	0,233
		<b>VEX</b>	0,145	0,032	61	4,46	<b>&lt; 0,0001</b>	0,081	0,209
		<b>CPRg</b>	0,200	0,086	61	2,31	<b>0,024</b>	0,029	0,370
		<b>STMg</b>	-0,067	0,257	61	-0,26	<b>0,794</b>	-0,574	0,440
Températures maximales du mois de septembre	<b>TXsep</b>	<b>CQP</b>	0,058	0,091	61	0,64	<b>0,525</b>	-0,121	0,236
		<b>VEX</b>	0,134	0,046	61	2,92	<b>0,005</b>	0,043	0,225
		<b>CPRg</b>	0,177	0,121	61	1,47	<b>0,147</b>	-0,061	0,415
		<b>STMg</b>	1,673	0,326	61	5,13	<b>&lt; 0,0001</b>	1,029	2,317

Températures maximales du mois d'octobre de l'année précédente	<b>TXsep_1</b>	<b>CQP</b>	-0,270	0,066	61	-4,09	<b>0,000</b>	-0,400	-0,139
		<b>VEX</b>	-0,177	0,033	61	-5,38	<b>&lt; 0,0001</b>	-0,242	-0,112
		<b>CPRg</b>	-0,065	0,094	61	-0,69	<b>0,491</b>	-0,251	0,120
		<b>STMg</b>	1,318	0,260	61	5,06	<b>&lt; 0,0001</b>	0,805	1,832
Températures maximales du mois d'octobre de l'année précédente	<b>TXoct_1</b>	<b>CQP</b>	0,150	0,053	61	2,84	<b>0,006</b>	0,046	0,255
		<b>VEX</b>	0,119	0,026	61	4,51	<b>&lt; 0,0001</b>	0,067	0,170
		<b>CPRg</b>	0,206	0,072	61	2,88	<b>0,006</b>	0,065	0,348
		<b>STMg</b>	-1,018	0,191	61	-5,34	<b>&lt; 0,0001</b>	-1,394	-0,642
Températures maximales du mois de novembre de l'année précédente	<b>TXnov_1</b>	<b>CQP</b>	0,385	0,052	61	7,40	<b>&lt; 0,0001</b>	0,283	0,488
		<b>VEX</b>	0,220	0,026	61	8,35	<b>&lt; 0,0001</b>	0,168	0,272
		<b>CPRg</b>	0,332	0,085	61	3,91	<b>0,000</b>	0,165	0,499
		<b>STMg</b>	-0,129	0,250	61	-0,52	<b>0,607</b>	-0,622	0,363
Températures maximales du mois de décembre de l'année précédente	<b>TXdec_1</b>	<b>CQP</b>	-0,367	0,068	61	-5,40	<b>&lt; 0,0001</b>	-0,502	-0,233
		<b>VEX</b>	-0,287	0,029	61	-9,92	<b>&lt; 0,0001</b>	-0,344	-0,230
		<b>CPRg</b>	-0,052	0,101	61	-0,51	<b>0,612</b>	-0,251	0,148
		<b>STMg</b>	0,120	0,299	61	0,40	<b>0,690</b>	-0,469	0,709
Températures moyennes du mois de janvier	<b>TYjan</b>	<b>CQP</b>	-0,048	0,038	61	-1,26	<b>0,211</b>	-0,123	0,027
		<b>VEX</b>	-0,013	0,020	61	-0,65	<b>0,521</b>	-0,052	0,026
		<b>CPRg</b>	0,065	0,051	61	1,29	<b>0,203</b>	-0,035	0,165
		<b>STMg</b>	-0,782	0,129	61	-6,05	<b>&lt; 0,0001</b>	-1,037	-0,527
Températures moyennes du mois de février	<b>TYfev</b>	<b>CQP</b>	0,161	0,033	61	4,91	<b>&lt; 0,0001</b>	0,097	0,226
		<b>VEX</b>	0,101	0,016	61	6,10	<b>&lt; 0,0001</b>	0,068	0,133
		<b>CPRg</b>	0,109	0,048	61	2,24	<b>0,029</b>	0,013	0,204
		<b>STMg</b>	-0,567	0,138	61	-4,10	<b>0,000</b>	-0,840	-0,295
Températures moyennes du mois de mars	<b>TYmar</b>	<b>CQP</b>	0,069	0,044	61	1,55	<b>0,126</b>	-0,019	0,156
		<b>VEX</b>	-0,015	0,024	61	-0,63	<b>0,528</b>	-0,061	0,032
		<b>CPRg</b>	0,017	0,060	61	0,28	<b>0,782</b>	-0,102	0,135
		<b>STMg</b>	-0,450	0,177	61	-2,54	<b>0,014</b>	-0,800	-0,100
Températures moyennes du mois d'avril	<b>TYavr</b>	<b>CQP</b>	0,266	0,041	61	6,44	<b>&lt; 0,0001</b>	0,185	0,348
		<b>VEX</b>	0,171	0,020	61	8,66	<b>&lt; 0,0001</b>	0,132	0,209
		<b>CPRg</b>	0,212	0,064	61	3,33	<b>0,001</b>	0,086	0,337
		<b>STMg</b>	-0,548	0,185	61	-2,95	<b>0,004</b>	-0,913	-0,182
Températures moyennes du mois de mai	<b>TYmai</b>	<b>CQP</b>	-0,253	0,068	61	-3,71	<b>0,000</b>	-0,387	-0,118
		<b>VEX</b>	-0,238	0,030	61	-7,94	<b>&lt; 0,0001</b>	-0,297	-0,179
		<b>CPRg</b>	-0,113	0,095	61	-1,18	<b>0,242</b>	-0,301	0,075
		<b>STMg</b>	-0,622	0,273	61	-2,27	<b>0,026</b>	-1,161	-0,083
Températures moyennes du mois de juin	<b>TYjun</b>	<b>CQP</b>	0,168	0,050	61	3,37	<b>0,001</b>	0,069	0,266
		<b>VEX</b>	0,177	0,022	61	7,95	<b>&lt; 0,0001</b>	0,133	0,221
		<b>CPRg</b>	0,175	0,069	61	2,55	<b>0,013</b>	0,039	0,310
		<b>STMg</b>	-0,192	0,203	61	-0,95	<b>0,348</b>	-0,593	0,209
Températures moyennes du mois de juillet	<b>TYjul</b>	<b>CQP</b>	0,320	0,140	61	2,29	<b>0,026</b>	0,044	0,595
		<b>VEX</b>	0,227	0,067	61	3,41	<b>0,001</b>	0,096	0,359
		<b>CPRg</b>	0,399	0,190	61	2,10	<b>0,040</b>	0,024	0,773
		<b>STMg</b>	-2,005	0,539	61	-3,72	<b>0,000</b>	-3,069	-0,942
Températures moyennes du mois d'août	<b>TYaou</b>	<b>CQP</b>	-0,030	0,076	61	-0,40	<b>0,691</b>	-0,179	0,119
		<b>VEX</b>	0,051	0,039	61	1,32	<b>0,192</b>	-0,026	0,128
		<b>CPRg</b>	0,086	0,101	61	0,86	<b>0,395</b>	-0,112	0,285
		<b>STMg</b>	-0,006	0,300	61	-0,02	<b>0,984</b>	-0,598	0,586

Températures moyennes du mois de septembre	<b>TYsep</b>	<b>CQP</b>	0,004	0,089	61	0,04	<b>0,969</b>	-0,173	0,180
		<b>VEX</b>	0,091	0,045	61	2,01	<b>0,049</b>	0,002	0,181
		<b>CPRg</b>	0,067	0,119	61	0,57	<b>0,574</b>	-0,167	0,301
		<b>STMg</b>	1,011	0,336	61	3,01	<b>0,004</b>	0,348	1,673
Températures moyennes du mois de septembre de l'année précédente	<b>TYsep_1</b>	<b>CQP</b>	-0,303	0,077	61	-3,93	<b>0,000</b>	-0,455	-0,151
		<b>VEX</b>	-0,226	0,037	61	-6,10	<b>&lt; 0,0001</b>	-0,299	-0,153
		<b>CPRg</b>	-0,174	0,108	61	-1,61	<b>0,112</b>	-0,388	0,039
		<b>STMg</b>	1,266	0,301	61	4,20	<b>&lt; 0,0001</b>	0,672	1,861
Températures moyennes du mois d'octobre de l'année précédente	<b>TYoct_1</b>	<b>CQP</b>	0,112	0,057	61	1,95	<b>0,056</b>	-0,002	0,225
		<b>VEX</b>	0,067	0,029	61	2,27	<b>0,027</b>	0,009	0,125
		<b>CPRg</b>	0,126	0,077	61	1,63	<b>0,109</b>	-0,027	0,278
		<b>STMg</b>	-1,259	0,201	61	-6,25	<b>&lt; 0,0001</b>	-1,656	-0,862
Températures moyennes du mois de novembre de l'année précédente	<b>TYnov_1</b>	<b>CQP</b>	0,416	0,061	61	6,81	<b>&lt; 0,0001</b>	0,295	0,536
		<b>VEX</b>	0,231	0,031	61	7,45	<b>&lt; 0,0001</b>	0,170	0,292
		<b>CPRg</b>	0,292	0,098	61	2,98	<b>0,004</b>	0,099	0,485
		<b>STMg</b>	0,041	0,286	61	0,14	<b>0,887</b>	-0,524	0,605
Températures moyennes du mois de décembre de l'année précédente	<b>TYdec_1</b>	<b>CQP</b>	-0,266	0,045	61	-5,89	<b>&lt; 0,0001</b>	-0,355	-0,177
		<b>VEX</b>	-0,206	0,019	61	-11,02	<b>&lt; 0,0001</b>	-0,243	-0,169
		<b>CPRg</b>	-0,101	0,068	61	-1,49	<b>0,140</b>	-0,235	0,032
		<b>STMg</b>	0,042	0,201	61	0,21	<b>0,835</b>	-0,354	0,438
Précipitations totales du mois de janvier	<b>Pjan</b>	<b>CQP</b>	0,020	0,004	61	4,64	<b>&lt; 0,0001</b>	0,011	0,028
		<b>VEX</b>	0,008	0,002	61	3,62	<b>0,001</b>	0,004	0,013
		<b>CPRg</b>	0,006	0,006	61	0,96	<b>0,340</b>	-0,006	0,018
		<b>STMg</b>	-0,043	0,018	61	-2,37	<b>0,021</b>	-0,079	-0,007
Précipitations totales du mois de février	<b>Pfev</b>	<b>CQP</b>	-0,012	0,007	61	-1,78	<b>0,080</b>	-0,024	0,001
		<b>VEX</b>	-0,011	0,003	61	-3,51	<b>0,001</b>	-0,018	-0,005
		<b>CPRg</b>	-0,004	0,009	61	-0,44	<b>0,659</b>	-0,021	0,013
		<b>STMg</b>	-0,029	0,026	61	-1,14	<b>0,260</b>	-0,079	0,021
Précipitations totales du mois de mars	<b>Pmar</b>	<b>CQP</b>	-0,018	0,003	61	-5,90	<b>&lt; 0,0001</b>	-0,023	-0,012
		<b>VEX</b>	-0,012	0,001	61	-8,44	<b>&lt; 0,0001</b>	-0,015	-0,009
		<b>CPRg</b>	-0,012	0,004	61	-2,64	<b>0,010</b>	-0,021	-0,003
		<b>STMg</b>	0,030	0,013	61	2,28	<b>0,026</b>	0,004	0,056
Précipitations totales du mois d'avril	<b>Pavr</b>	<b>CQP</b>	0,008	0,002	61	4,57	<b>&lt; 0,0001</b>	0,005	0,012
		<b>VEX</b>	0,007	0,001	61	9,39	<b>&lt; 0,0001</b>	0,006	0,009
		<b>CPRg</b>	0,005	0,003	61	2,04	<b>0,045</b>	0,000	0,010
		<b>STMg</b>	-0,006	0,008	61	-0,74	<b>0,462</b>	-0,020	0,009
Précipitations totales du mois de mai	<b>Pmai</b>	<b>CQP</b>	-0,013	0,002	61	-7,19	<b>&lt; 0,0001</b>	-0,017	-0,010
		<b>VEX</b>	-0,009	0,001	61	-10,84	<b>&lt; 0,0001</b>	-0,011	-0,007
		<b>CPRg</b>	-0,007	0,003	61	-2,46	<b>0,017</b>	-0,013	-0,001
		<b>STMg</b>	0,014	0,009	61	1,57	<b>0,121</b>	-0,004	0,031
Précipitations totales du mois de juin	<b>Pjun</b>	<b>CQP</b>	0,004	0,002	61	1,97	<b>0,054</b>	0,000	0,008
		<b>VEX</b>	0,003	0,001	61	2,58	<b>0,012</b>	0,001	0,005
		<b>CPRg</b>	0,003	0,003	61	1,04	<b>0,302</b>	-0,003	0,008
		<b>STMg</b>	-0,037	0,008	61	-4,93	<b>&lt; 0,0001</b>	-0,052	-0,022
Précipitations totales du mois de juillet	<b>Pjul</b>	<b>CQP</b>	0,005	0,002	61	2,04	<b>0,046</b>	0,000	0,010
		<b>VEX</b>	0,003	0,001	61	2,67	<b>0,010</b>	0,001	0,006
		<b>CPRg</b>	0,004	0,003	61	1,25	<b>0,218</b>	-0,002	0,011
		<b>STMg</b>	-0,004	0,010	61	-0,43	<b>0,671</b>	-0,024	0,015



Précipitations totales du mois d'août	<b>Paou</b>	<b>CQP</b>	0,010	0,002	61	3,91	<b>0,000</b>	0,005	0,015
		<b>VEX</b>	0,009	0,001	61	7,70	<b>&lt; 0,0001</b>	0,006	0,011
		<b>CPRg</b>	0,009	0,004	61	2,44	<b>0,018</b>	0,002	0,015
		<b>STMg</b>	-0,030	0,010	61	-2,89	<b>0,005</b>	-0,050	-0,009
Précipitations totales du mois de septembre	<b>Psep</b>	<b>CQP</b>	0,007	0,003	61	2,44	<b>0,018</b>	0,001	0,013
		<b>VEX</b>	0,008	0,001	61	5,75	<b>&lt; 0,0001</b>	0,005	0,011
		<b>CPRg</b>	0,005	0,004	61	1,34	<b>0,187</b>	-0,003	0,014
		<b>STMg</b>	-0,032	0,012	61	-2,67	<b>0,010</b>	-0,055	-0,008
Précipitations totales du mois de septembre de l'année précédente	<b>Psep_1</b>	<b>CQP</b>	-0,006	0,005	61	-1,19	<b>0,240</b>	-0,015	0,004
		<b>VEX</b>	-0,005	0,002	61	-2,34	<b>0,023</b>	-0,010	-0,001
		<b>CPRg</b>	-0,008	0,006	61	-1,26	<b>0,213</b>	-0,020	0,005
		<b>STMg</b>	-0,016	0,018	61	-0,89	<b>0,379</b>	-0,053	0,020
Précipitations totales du mois d'octobre de l'année précédente	<b>Poct_1</b>	<b>CQP</b>	0,000	0,002	61	-0,24	<b>0,813</b>	-0,004	0,003
		<b>VEX</b>	-0,003	0,001	61	-3,14	<b>0,003</b>	-0,004	-0,001
		<b>CPRg</b>	-0,003	0,002	61	-1,25	<b>0,215</b>	-0,007	0,002
		<b>STMg</b>	0,010	0,007	61	1,45	<b>0,153</b>	-0,003	0,023
Précipitations totales du mois de novembre de l'année précédente	<b>Pnov_1</b>	<b>CQP</b>	0,004	0,002	61	1,72	<b>0,091</b>	-0,001	0,008
		<b>VEX</b>	0,002	0,001	61	1,47	<b>0,148</b>	-0,001	0,004
		<b>CPRg</b>	-0,002	0,003	61	-0,84	<b>0,402</b>	-0,008	0,003
		<b>STMg</b>	0,039	0,007	61	5,36	<b>&lt; 0,000 1</b>	0,025	0,053
Précipitations totales du mois de décembre de l'année précédente	<b>Pdec_1</b>	<b>CQP</b>	0,005	0,001	61	4,52	<b>&lt; 0,000 1</b>	0,003	0,007
		<b>VEX</b>	0,003	0,001	61	4,22	<b>&lt; 0,0001</b>	0,001	0,004
		<b>CPRg</b>	0,002	0,002	61	1,17	<b>0,245</b>	-0,001	0,005
		<b>STMg</b>	-0,014	0,005	61	-2,83	<b>0,006</b>	-0,024	-0,004

## Phase de modélisation – étape 4 – régression multivariée (paragraphe 5.2.4)

Tableau A2 : Résultat des régressions univariées des termes quadratiques et d'interaction pour le groupe *Culex pipiens*-restuans.

CPRg		
Variable	t	p-value
Agri	5,5	< 0,0001
Prairie	-2,014	0,048
Boise	-3,881	0
TNsep_1	-2,318	0,024
Agri x Prairie		0,779
Agri x Boise		0,674
Prairie x Boise	-2,445	0,017
Agri x Agri	4,5	< 0,0001
Agri x TNsep_1		0,11
Prairie x TNsep_1		0,326
Boise x TNsep_1	3 412	0,001

Tableau A3 : Résultat des régressions univariées des termes quadratiques et d'interaction pour *Coquillettidia perturbans*.

CQP		
Variable	t	p-value
Agri	2,349	0,022
Urbain	-5,35	< 0,0001
Boise	4,18	< 0,0001
Pjan	4,637	< 0,0001
Psep	2,439	0,018
Pjul	2,037	0,046
Paou	3,915	0,000
Agri x Urbain		0,717
Agri x Boise		0,161
Agri x Pjan		0,492
Agri x Psep		0,557
Urbain x Boise	-2,742	0,008
Urbain x Pjan		0,255
Urbain x Psep		0,62
Boise x Pjan		0,056
Boise x Psep		0,869
Pjan x Psep		0,937
Agri x Pjul		0,41
Urbain x Pjul		0,316
Boise x Pjul		0,343
Pjan x Pjul		0,762
Psep x Pjul		0,141
Agri x Paou		0,411
Urbain x Paou		0,596
BOIS x Paou	-0,988	0,327
Pjan x Paou	0,858	0,394
Psep x Paou	3,263	0,002
Pjul x Paou	-3,021	0,004

Tableau A4 : Résultat des régressions univariées des termes quadratiques et d'interaction pour le groupe Ochlerotatus stimulans.

STMg		
Variable	t	p-value
Urbain	-6,765	< 0,0001
Prairie	2,215	0,031
Boise	6,228	< 0,0001
Pjan1km	-2,369	0,021
Pmar1km	2,278	0,026
Pjun1km	-4,934	< 0,0001
Paou1km	-2,89	0,005
Psep1km	-2,671	0,010
Urbain x Prairie		0,591
Urbain x Boise		0,076
Prairie x Boise		0,436
Boise x Boise	2,599	0,012
Urbain x Pjan		0,798
Prairie x Pjan		0,685
Boise x Pjan		0,434
Urbain x Pmar		0,633
Prairie x Pmar		0,922
Boise x Pmar		0,265
Pjan x Pmar		0,884
Urbain x Pjun		0,812
Prairie x Pjun		0,431
Boise x Pjun		0,824
Pjan x Pjun		0,297
Urbain x Paou		0,535
Prairie x Paou		0,803
Boise x Paou		0,248
Pjan x Paou		0,394
Urbain x Psep		0,559
Prairie x Psep		0,772
Boise x Psep		0,141
Pjan x Psep		0,765
Pmar x Pjun		0,717
Pmar x Paou		0,514
Pmar x Psep		0,207
Pjun x Paou	3,736	0
Pjun x Psep		0,149
Paou x Psep	-3,518	0,001

Tableau A5 : Résultat des régressions univariées des termes quadratiques et d'interaction pour Aedes vexans.

VEX		
Variable	t	p-value
Eau	-2,906	0,005
Agri	4,573	< 0,0001
Pfev	-3,511	0,001
Psep_1	-2,341	0,023
Pjan	3,617	0,001
Pjul	2,67	0,010
Psep	5,75	< 0,0001
Eau x Agri		0,165
Eau x Pfev		0,329
Agri x Agri	3,118	0,003
Eau x Psep_1		0,155
Agri x Pfev		0,492
Agri x Psep_1	-2,457	0,017
Pfev x Psep_1		0,55
Eau x Pjan		0,4
Agri x Pjan		0,481
Pfev x Pjan		0,62
Psep_1 x Pjan	-2,303	0,025
Eau x Pjul		0,828
Agri x Pjul	2,243	0,029
Pfev x Pjul		0,516
Psep_1 x Pjul		0,076
Eau x Psep		0,215
Agri x Psep		0,068
Pfev x Psep		0,622
Psep_1 x Psep		0,435
Pjan x Pjul		0,811
Pjan x Psep	4,096	0
Pjul x Psep		0,381